X-643-69-368

NASA TM X- 63784

INTERPOLATION FOR FUNCTIONS OF SEVERAL VARIABLES

JOEL S. COHEN

AUGUST 1969

GSFC

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

X-643-69-368

INTERPOLATION FOR FUNCTIONS OF SEVERAL VARIABLES

Joel S. Cohen

August 1969

GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland

CONTENTS

INTERPOLATION FOR FUNCTIONS OF SEVERAL VARIABLES

CHAPTER I

GENERAL INTERPOLATION

1.1 INTRODUCTION

The usual method for interpolation of functions of several variables has been to iterate the interpolation process for each variable taken separately. By its very nature, this process has forced the base points to be located at the corners of a rectangular or triangular grid. These are, however, generalizations of some of the classical interpolation formulas to much larger classes of points. This report is a summary of some of the algebraic techniques in interpolation theory with particular emphasis on the application of these techniques to functions of several variables. In particular, we discuss what appears to be the most natural generalizations of the familiar Lagrange formula, Newton divided difference scheme and Aitken interpolation scheme to functions of several variables. We shall show that both the Lagrange formula and Newton formula can be derived from more general theorems on finite dimensional vector spaces. When applied to functions of several variables, these formulae are applicable to arbitrarily located base points as long as the determinants appearing in the denominators are not zero. On the other hand, the Aitken scheme does not have a generalization to arbitrary finite dimensional vector spaces. In this case we present an iteration scheme, due to Thacher and Milne [19], which is valid for certain restricted classes of points.

For the single variable case, the purpose of the various interpolation schemes has been to put the interpolation polynomial in a form which is more convenient for numerical computation. In the several variable case this is only accomplished to a much smaller degree. For the single variable case, the Vandermonde determinant plays a fundamental role in the derivation of the various interpolation schemes and in almost every case the simple single variable formulae can be derived from the general theorems on vector spaces by using the evaluation properties of this determinant. In the several variable case this tool is not available; for this reason the formulae will be considerably more complicated.

For a more complete discussion of the topics discussed in this report we refer the reader to the list of papers in the bibliography. Almost all the material presented here can be found in one form or another in the work of Milne, Arntzen, Reynolds, Wheelock ([8]); Thacher ([16], [17], [18]); Milne and Thacher ([19]); Salzer ([11]); and Curry ([1]).

Preliminaries and Notations

We shall assume the reader is familiar with the basic single variable interpolation techniques as found in Hildebrand [5] or Householder [7]. In addition we assume the reader is familiar with the basic techniques of linear algebra as found in Hoffman and Kunze [22]. We shall briefly recall a few definitions and results from linear algebra and matrix theory which will be needed in the sections to follow.

Let V and W be finite dimensional vector spaces. A linear operator T is a function

$$T : V \longrightarrow W$$

which satisfies the property

$$T(c_1 u + c_2 v) = c_1 (T(u)) + c_2 (T(v)) ,$$

where c_1 and c_2 are real numbers and u and v belong to V. When the range space W is the vector space of real numbers it is customary to use the term linear functional. One can easily show the collection of linear functionals defined in V is itself a vector space. This vector space is known as the dual space for V and is usually denoted by V^*. The following results can be found in Hoffman and Kunze ([22], p. 90-96).

Theorem (1.1): Let V be an n dimensional vector space. Then

1) $\dim V = \dim V^*$;

2) If $v_1, \ldots, v_n$ is a basis for V, then, there exists a unique basis $L_1, L_2, \ldots, L_n$ for V^* which satisfies the property

$$L_i (v_j) = \delta_{ij} .$$

This basis is known as the dual basis associated with $v_1, \ldots, v_n$.

3) $V = (V^*)^*$.

We shall use the following notation:

1) The symbol R will denote the real numbers; R^n will denote the usual n dimensional Euclidean space.

2) If $u_1, \ldots, u_p$ are members of a vector space V, then

$$\mathrm{sp}\{u_1, \cdots, u_p\}$$

will denote the vector space spanned by these elements.

3) If $[a_{ij}]$ is an $n \times n$ matrix, the determinant of $[a_{ij}]$ will be denoted by

$$\det[a_{ij}] .$$

Let $A_{i_0 j_0}$ be the matrix obtained from $[a_{ij}]$ by removing the i_0^{th} row and j_0^{th} column. Then we shall use the notation

$$\mathrm{Cof}\left(a_{i_0 j_0}\right) = (-1)^{i_0 + j_0} \det A_{i_0 j_0} .$$

One particular matrix occurs quite often in interpolation theory. If x_0, $x_1, \ldots, x_n$ are real numbers, we can consider the matrix

$$\begin{bmatrix} 1 & x_0 & x_0^2 \cdots & x_0^n \\ 1 & x_1 & x_1^2 \cdots & x_1^n \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & x_n & x_n^2 \cdots & x_n^n \end{bmatrix} .$$

The determinant of this matrix is known as the Vandermonde determinant. It can be evaluated by the following formula:

$$\det \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix} = \prod_{i>j} (x_i - x_j)$$

([7], p. 91).

1.2 GENERAL INTERPOLATION

In interpolation theory we attempt to approximate a function f defined in a region G of R^n by a finite linear combination

$$\phi(\overline{x}) = \sum_{j=0}^{m} a_j \phi_j (\overline{x}) , \tag{1.2-1}$$

where $\{\phi_j (\overline{x})\}$ is a collection of linearly independent functions defined in G. In the most restricted form of interpolation we determine the constants $\{a_j\}$ by choosing a collection of base points $\overline{x}_0, \overline{x}_1, \ldots, \overline{x}_m$ in G and requiring that

$$\phi(\overline{x}_i) = f(\overline{x}_i) , \tag{1.2-2}$$

for $i = 0, 1, 2, \ldots, m$. Applying this condition to Equation (1.2-1) we obtain the $m + 1$ equations

$$f(\overline{x}_i) = \sum_{j=0}^{m} a_j \phi_j (\overline{x}_i) \tag{1.2-3}$$

The condition

$$D = \det\left[\phi_j (\overline{x}_i)\right] \neq 0 \tag{1.2-4}$$

is necessary and sufficient to guarantee that the system (1.2-3) has a unique solution. Therefore when (1.2-4) is satisfied, we obtain

$$\phi(\bar{x}) = \sum_{j=0}^{m} \frac{D_j}{D} \phi_j(\bar{x}) . \tag{1.2-5}$$

where D_j is the determinant of the matrix obtained by replacing the j^{th} column of the matrix $\left[\phi_j\left(x_i\right)\right]$ by the column vector $\left[f\left(x_i\right)\right]$.

Alternately, some authors ([19], [7], p. 186) define $\phi(\bar{x})$ by the formula

$$\Delta = \det\begin{pmatrix} \phi_0\left(\bar{x}_0\right)\cdots & \phi_m\left(\bar{x}_0\right) & f\left(\bar{x}_0\right) \\ \vdots & \vdots & \vdots \\ \phi_0\left(\bar{x}_m\right)\cdots & \phi_m\left(\bar{x}_m\right) & f\left(\bar{x}_m\right) \\ \phi_0(\bar{x}) \cdots & \phi_m(\bar{x}) & \phi(\bar{x}) \end{pmatrix} = 0 \tag{1.2-6}$$

If $D \neq 0$, formula (1.2-6) defines $\phi(\bar{x})$ uniquely as a linear combination of ϕ_0, $\phi_1, \ldots, \phi_m$. In order to show $\phi(\bar{x})$ defined in this way satisfies the condition imposed by Equation (1.2-2), we set $\bar{x} = \bar{x}_j$ and expand (1.2-6) by minors of the last column; we obtain

$$\sum_{i=0}^{m} f\left(\bar{x}_i\right) \operatorname{Cof}\left(f\left(\bar{x}_i\right)\right) + D\phi\left(\bar{x}_j\right) = 0 \tag{1.2-7}$$

However, since $\bar{x} = \bar{x}_j$ we have

$$\operatorname{Cof}\left(f\left(\bar{x}_i\right)\right) = \begin{cases} 0 & \text{if} \quad i \neq j \\ -D & \text{if} \quad i = j \end{cases} .$$

Therefore Equation (1.2-7) becomes

$$-D\phi(\overline{x}_j) = -D f(\overline{x}_j)$$

or

$$\phi(\overline{x}_j) = f(\overline{x}_j) .$$

We summarize the preceding discussion in the following theorem:

Theorem (1.2.1): Let $\phi_0, \phi_1, \ldots, \phi_m$ be a collection of linearly independent functions defined in a region G of R^n and let $\overline{x}_0, \overline{x}_1, \ldots, \overline{x}_m$ be $m+1$ points in G. Suppose f is an arbitrary function defined in G. Then, there is a unique function $\phi(\overline{x})$, which is a linear combination of $\{\phi_j(\overline{x})\}$ and which satisfies the condition

$$f(\overline{x}_i) = \phi(\overline{x}_i) ,$$

if and only if

$$D = \det\left[\phi_j(\overline{x}_i)\right] \neq 0 .$$

If this condition is satisfied, $\phi(\overline{x})$ is given by (1.2-5) or (1.2-6).

Definition (1.2.2): We refer to the interpolation problem considered in Theorem (1.2.1) as the restricted interpolation problem. The unique function $\phi(\overline{x})$ is known as the interpolating function for f at the points $\overline{x}_0, \overline{x}_1, \ldots, \overline{x}_m$.

We relate the preceding discussion to the problem of polynomial interpolation for functions of several real variables.

Example (1.2.3): Let $f(x)$ be a real valued function defined in an interval $[a, b]$ and let $x_0, x_1, \ldots, x_m$ be $m+1$ distinct points in $[a, b]$. For our basis

functions we take the collection

$$
\begin{aligned}
\phi_0(x) &= 1 \\
\phi_1(x) &= x \\
&\vdots \\
\phi_m(x) &= x^m .
\end{aligned}
$$

In this case, the determinant in (1.2-4) takes the form

$$
D = \det\begin{pmatrix} 1 & x_0 & \cdots & x_0^m \\ 1 & x_1 & \cdots & x_1^m \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^m \end{pmatrix}.
$$

This determinant is the Vandermonde determinant and its value is given by

$$
D = \det\begin{pmatrix} 1 & x_0 & \cdots & x_0^m \\ 1 & x_1 & \cdots & x_1^m \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^m \end{pmatrix} = \prod_{i<j} (x_i - x_j) . \qquad (1.2\text{-}8)
$$

Since the points $\{x_i\}$ are distinct, this determinant is not zero and by Theorem (1.2.1) there is a polynomial

$$
p(x) = \sum_{i=0}^{m} a_i x^i
$$

such that

$$p(x_i) = f(x_i) \qquad i = 0, 1, 2, \cdots, m.$$

<u>Example (1.2.4)</u>: Let G be a region of R^2 containing the points

$$\bar{x}_0 = (-2, 0),$$

$$\bar{x}_1 = \left(-3, -\frac{1}{2}\right),$$

$$\bar{x}_2 = (0, -2),$$

$$\bar{x}_3 = \left(1, -\frac{3}{2}\right)$$

Let $f(\bar{x})$ be a function defined in G which takes the values

$$f\big((-2, 0)\big) = 4,$$

$$f\left(\left(-3, -\frac{1}{2}\right)\right) = 5,$$

$$f\big((0, -2)\big) = -10,$$

$$f\left(\left(1, -\frac{3}{2}\right)\right) = -5$$

For our basis functions we take the polynomials

$$\phi_0(x, y) = 1,$$

$$\phi_1(x, y) = x,$$

$$\phi_2(x, y) = y ,$$

$$\phi_3(x, y) = xy .$$

In this case, Equation (1.2-4) becomes

$$D = \det\left[\phi_j\left(\bar{x}_i\right)\right] = \det\begin{pmatrix} 1 & -2 & 0 & 0 \\ 1 & -3 & -\frac{1}{2} & \frac{3}{2} \\ 1 & 0 & -2 & 0 \\ 1 & 1 & -\frac{3}{2} & -\frac{3}{2} \end{pmatrix} = 0 . \quad (1.2\text{-}9)$$

Since the above determinant is zero, the conditions of Theorem (1.2.1) are not satisfied and we cannot produce a unique polynomial in the form

$$\phi(x, y) = a_0 + a_1 x + a_2 y + a_3 xy$$

which satisfies the condition

$$\phi\left(\bar{x}_i\right) = f\left(\bar{x}_i\right)$$

for $i = 0, 1, 2, 3$.

Example (1.2.5): In this example we consider the same base points and function f considered in Example (1.2.4). However, in this example we define

$$\phi_0(x, y) = 1 ,$$

$$\phi_1(x, y) = x ,$$

$$\phi_2(x, y) = y ,$$

$$\phi_3(x, y) = x^2 .$$

In this case the determinant in Equation (1.2-4) becomes

$$\det \begin{pmatrix} 1 & -2 & 0 & 4 \\ 1 & -3 & -\frac{1}{2} & 9 \\ 1 & 0 & -2 & 0 \\ 1 & 1 & -\frac{3}{2} & 1 \end{pmatrix} = 18 \neq 0 .$$

Therefore, this time the conditions for Theorem (2.1.1) are satisfied. Using Equation (1.2-5) we obtain

$$\phi(x, y) = x^2 + 2x + 6y + 2 .$$

Remark (1.2.6): The above examples point out a fundamental difference between single variable interpolation and interpolation for functions of several variables. In the single variable case we are always guaranteed a unique solution to the restricted interpolation problem by choosing the functions $1, x, x^2, \ldots,$ x^m for our basis functions. However, in the several variables case it may not be at all clear which monomials we must use in order to guarantee a unique solution. Except for certain special cases, the author is not familiar with any "short cut" techniques for determining whether or not a unique solution does exist. In most cases one must evaluate the determinant in Equation (1.2-4).

There is, however, a geometric approach to this problem which might be valuable in some special cases. In Example (1.2.4) we can make the observations that the four base points lie in the algebraic curve

$$p(x, y) = y + xy + x + 2 = 0 .$$

This observation is in fact equivalent to the singularity of the matrix in (1.2-9). We relate the two ideas in the following theorem.

Theorem (1.2.7): Let $\phi_0, \phi_1, \ldots, \phi_m$ be $m+1$ linearly independent functions defined in a region G of R^n and let $\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_m$ be $m+1$ points in G. Then

$$\det\left[\phi_j(\bar{x}_i)\right] = 0$$

if and only if there exists a nontrivial algebraic curve in R^n in the form

$$p(\bar{x}) = \sum_{i=0}^{m} a_i \phi_i(\bar{x}) = 0$$

such that $p(\bar{x}_i) = 0$, for $i = 0, 1, \ldots, m$.

Proof: Suppose the points $\bar{x}_0, \ldots, \bar{x}_m$ lie on the curve

$$p(\bar{x}) = \sum_{i=0}^{m} a_i \phi_i(\bar{x}) = 0,$$

where not all the coefficients $\{a_i\}$ are zero. Substituting the points $x_0, \ldots, x_m$ in this equation we get the homogeneous system

$$\begin{aligned} \sum_{i=0}^{m} a_i \phi_i(\bar{x}_0) &= 0 \\ \sum_{i=0}^{m} a_i \phi_i(\bar{x}_1) &= 0 \\ &\;\;\vdots \\ \sum_{i=0}^{m} a_i \phi_i(\bar{x}_m) &= 0. \end{aligned} \tag{1.2-10}$$

Since not all the coefficients a_i are zero, the above system has a nontrivial solution. It follows that

$$\det\left[\phi_j\left(\bar{x}_i\right)\right] = 0$$

Conversely, if the above determinant is zero the homogeneous system (1.2.10) must have at least one nontrivial solution $a_0, a_1, \ldots, a_m$. The curve

$$p(\bar{x}) = \sum_{i=0}^{m} a_i \phi_i (\bar{x}) = 0$$

satisfies the condition $p\left(\bar{x}_i\right) = 0$ for $i = 0, 1, 2, \ldots, m$.

∎

1.3 INTERPOLATION ON TWO DIMENSIONAL GRIDS

In most practical problems Theorem (1.2.5) is of little value. However, for the case of restricted polynomial interpolation we can apply Theorem (1.2.7) to rectangular and triangular grids of points in order to obtain a sufficient collection of monomials to guarantee a unique interpolation polynomial.

We shall use the following lemma:

Lemma (1.3.1): Let $p(x, y)$ be a polynomial in two variables and suppose for a fixed real number x_0 the polynomial $p(x_0, y) = 0$ for all y. Then there exists a polynomial $g(x, y)$ such that

$$p(x, y) = \left(x - x_0\right) g(x, y)$$

Proof: We can write $p(x, y)$ in the form

$$p(x, y) = \sum_{i=0}^{n} p_i (y) x^i , \tag{1.3-1}$$

where each coefficient $p_i(y)$ is a polynomial in y. Since for each fixed y, the polynomial $p(x_0, y)$ is equal to zero, we can write

$$p(x, y) = (x - x_0) \sum_{i=0}^{n-1} b_i(y) x^i . \tag{1.3-2}$$

We must show the coefficients $b_i(y)$ are polynomials in y. However, comparing the coefficients of x^i in Equations (1.3-1) and (1.3-2), we obtain

$$b_{n-1}(y) = p_n(y)$$

and

$$b_{n-j}(y) = p_{n-j+1}(y) + x_0 b_{n-j+1}$$

for $j = 2, 3, \ldots, n$. Therefore, the coefficients $b_i(y)$ are polynomials in the variable y.

■

We now turn our attention to rectangular grids in R^2. Let (x_i, y_j), $(i = 0, \ldots, n)$, $(j = 0, 1, \ldots, m)$ be $(n + 1)(m + 1)$ points on a rectangular lattice in R^2 (see Figure (1.3.1)).

Let F be a function defined in a region G of R^2 containing the above rectangular lattice. In the following theorem we obtain a collection of monomials which is sufficient to guarantee a unique solution to the restricted interpolation problem on the above grid.

Theorem (1.3.2): Let f be a function defined in region G of R^2 containing the rectangular lattice in Figure (1.3.1). Then, there exists a unique interpolating polynomial $\phi(x, y)$ in the form

$$\phi(x, y) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_{ij} x^i y^j , \tag{1.3-3}$$

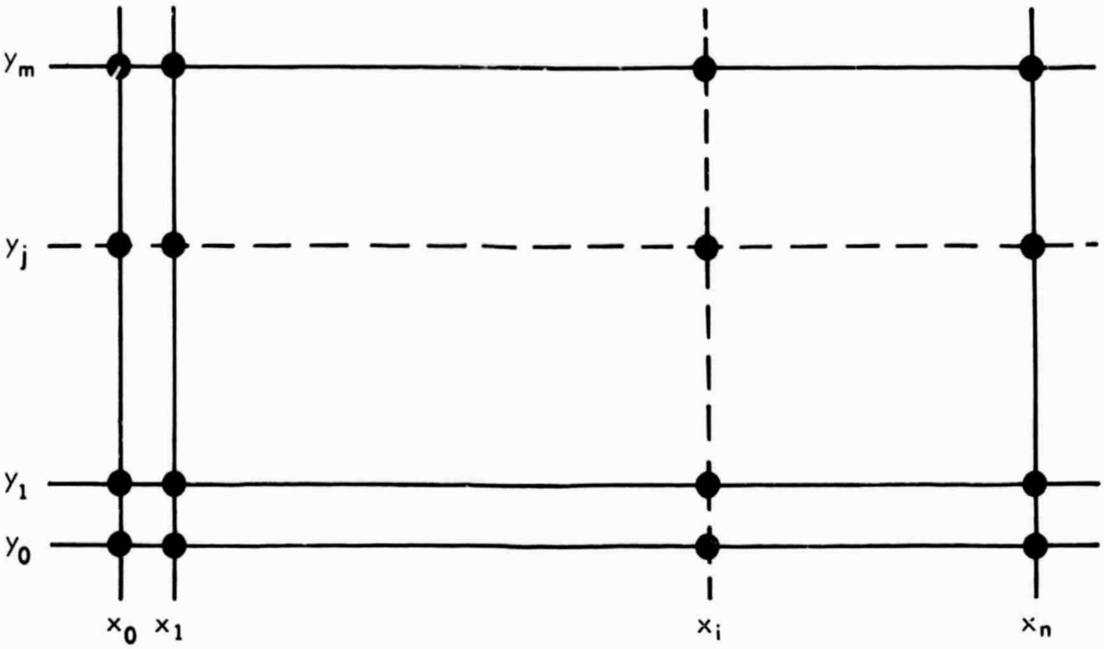

Figure (1.3.1)

such that

$$\phi(x_i, y_j) = f(x_i, y_j) ,$$

for $(i = 0, 1, \ldots, n)$, $(j = 0, 1, \ldots, m)$.

<u>Proof:</u> We prove this theorem by contradiction. Suppose a unique solution in the form of Equation (1.3-3) did not exist. Then, by Theorem (1.2.7), there exists a nondegenerate algebraic curve

$$p_0(x, y) = \sum_{i,j}^{m,n} b_{ij} x^i j^j = 0$$

such that

$$p_0(x_i, y_j) = 0$$

for $i = 0, 1, \ldots, n$ and $j = 0, 1, 2, \ldots, m$. Let $x = x_0$. The polynomial $p_0(x_0, y)$ has degree less than or equal to m in the variable y. Since it vanishes at the $m+1$ points $y_0, y_1, \ldots, y_m$ it must vanish for all y. By Lemma (1.3.1), there exists a polynomial $p_1(x, y)$ such that

$$p_0(x, y) = \left(x - x_0\right) p_1(x, y).$$

We can apply the above process for the point x_1 and the polynomial $p_1(x, y)$. Since $p_1(x_1, y)$ is a polynomial in y of degree less than or equal to m which vanishes at $y_0, y_1, \ldots, y_m$, it must vanish for all y. Therefore, by Lemma (1.3.1), there exists a polynomial $p_2(x, y)$ such that

$$\begin{aligned} p_0(x, y) &= \left(x - x_0\right) p_1(x, y) \\ &= \left(x - x_0\right)\left(x - x_1\right) p_2(x, y). \end{aligned}$$

Continuing in this manner, we obtain a polynomial $p_n(x, y)$ such that

$$p_0(x, y) = \prod_{i=0}^{n-1} \left(x - x_i\right) p_{n-1}(x, y). \tag{1.3-4}$$

Since $p_0(x, y)$ has degree less than or equal to n in the variable x, the polynomial $p_{n-1}(x, y)$ can not depend on x. Therefore, we can write

$$p_0(x, y) = \prod_{i=0}^{n-1} \left(x - x_i\right) p_{n-1}(y).$$

Now let $x = x_n$. Since the polynomial $p_0(x_n, y)$ has degree at most m in the variable y and since $p_0(x_n, y)$ is zero for $y = y_0, y_1, \ldots, y_m$, it follows that

$p_0(x_n, y) = 0$ for all y. Furthermore, since

$$\prod_{i=0}^{n-1} (x_n - x_i) \neq 0$$

and since $p_n(y)$ does not depend on x, it follows that $p_n(y)$ is identically equal to zero. But then $p_0(x, y)$ is identically equal to zero which contradicts our assumption that $p_0(x, y)$ be nondegenerate. ■

The previous theorem does not guarantee a polynomial of lowest possible total degree. For example, consider the grid in Figure 1.3.2. If f is any function

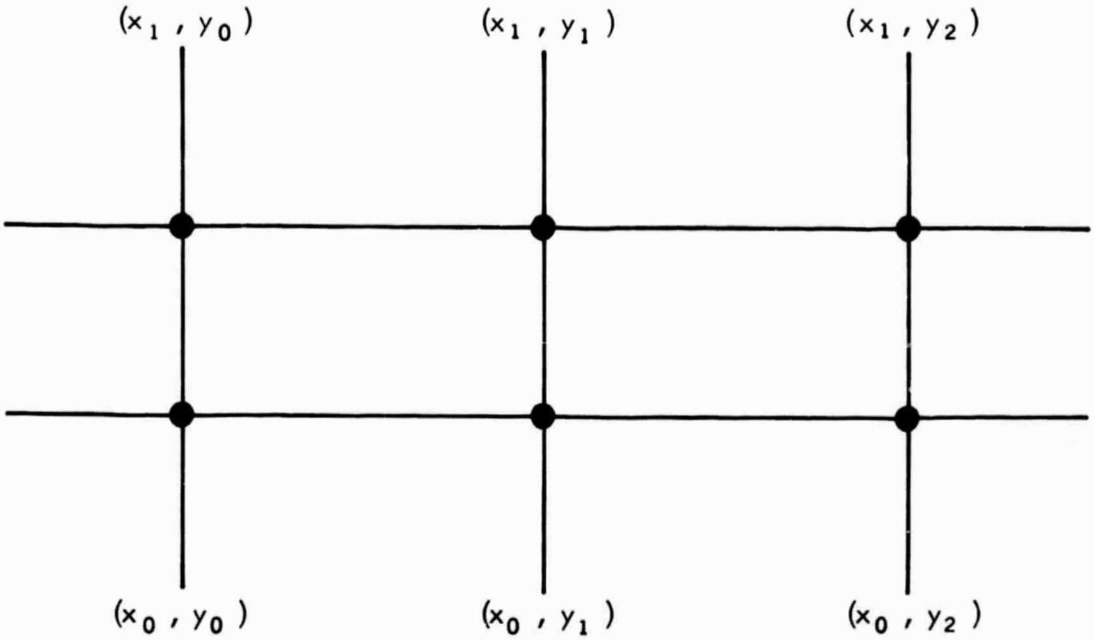

Figure (1.3.2)

defined in a region G containing the above grid, then the previous theorem guarantees an interpolation polynomial of the form

$$\phi(x, y) = a_0 + a_1 x + a_2 x^2 + a_3 y + a_4 xy + a_5 x^2 y$$

such that $\phi(x, y)$ and $f(x, y)$ agree at the points on the above grid. The polynomial $\phi(x, y)$ has total degree 3. There are, however, six monomials $1, x, y, x^2, xy, y^2$ with total degree at most 2; unfortunately unique interpolation with these monomials is impossible. One can easily check that the points in Figure (1.3.2) lie on the nondegenerate algebraic curve

$$p(x, y) = (y - y_0)(y - y_1) = 0 .$$

By Theorem (1.2.7), interpolation with the minimals $1, x, y, x^2, xy, y^2$ is impossible.

The following theorem shows the above problem does not occur on triangular grids. By interpolating with the base points on a triangular grid we can obtain an interpolating polynomial of lowest possible degree.

Theorem (1.3.3): Let f be a function defined in a region G containing the triangular grid in Figure (1.3.3). Then, there exists a polynomial of total

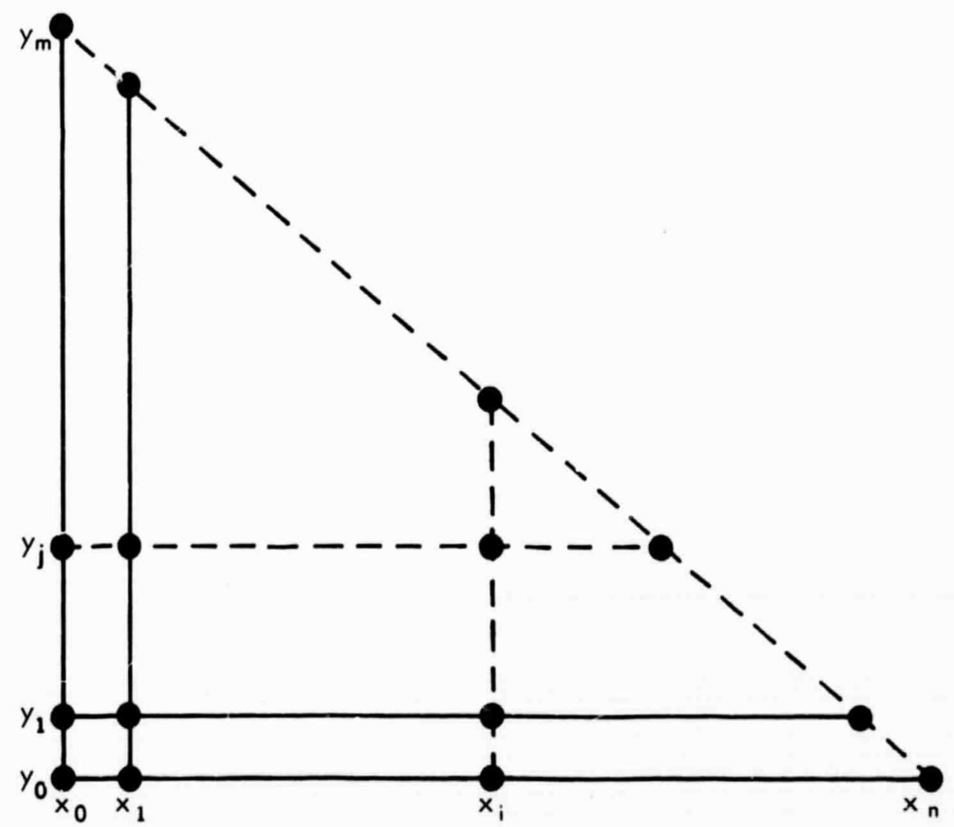

Figure (1.3.3)

degree n such that f and ϕ agree at the points of the above triangular lattice.

Proof: We shall prove the theorem by contradiction: Suppose a unique polynomial $\phi(x, y)$ of total degree n did not exist. Then, by Theorem (1.2.7), there exists an algebraic curve $p(x, y)$ of total degree n which vanishes at the points on the above triangular lattice. Let us write $p(x, y)$ in the form

$$p(x, y) = \sum_{i=0}^{n} A_i(y)\, x^i$$

where $A_i(y)$ has degree $n - i$ in the variable y. If we set $y = y_0$, then

$$p(x, y_0) = \sum_{i=0}^{n} A_i(y_0)\, x^i$$

is a polynomial in the variable x which vanishes at the $n + 1$ points $x_0, x_1, \ldots, x_n$. Therefore, since $p(x, y_0)$ has degree n, it follows that $p(x, y_0)$ is zero for all values of x. In particular, the coefficient $A_n(y)$ which, as a polynomial in y, has degree zero must vanish at $y = y_0$. It follows that $A_n(y)$ is identically zero. Next we set $y = y_1$. In this case, the polynomial $p(x, y_1)$ will vanish at the n points $x_0, x_1, \ldots, x_{n-1}$. Therefore, since $p(x, y_1)$ has degree $n - 1$ in the variable x, it must be identically zero. Now we recall that $A_{n-1}(y)$ has the form

$$A_{n-1}(y) = cy + d .$$

However, from the above remarks we have

$$A_{n-1}(y_0) = 0$$

$$A_{n-1}(y_1) = 0 ,$$

which implies A_{n-1} is identically zero. In a similar way, one can show $A_{n-2}(y)$, $A_{n-3}(y), \ldots, A_0(y)$ are all identically zero. Therefore, $p(x, y)$ is identically zero which contradicts the fact that $p(x, y)$ be nondegenerate.

$\blacksquare$

The above techniques can be used to find monomials sufficient to guarantee a unique solution for points arbitrarily located at certain vertices of a rectangular grid. We illustrate the technique in the following example.

Example (1.3.4): Consider the problem of finding a sufficient set of monomials which will guarantee a unique interpolating polynomial for functions defined in a region containing the grid in Figure (1.3.4).

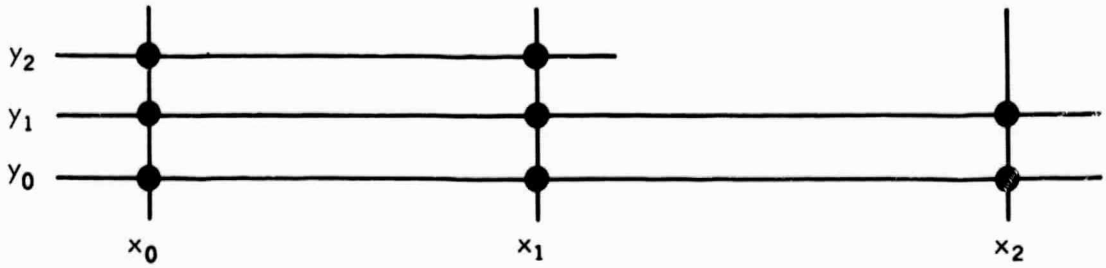

Figure (1.3.4)

We shall show that the monomials

$$\begin{array}{lll} y^2, & y^2x, & \\ y, & yx, & yx^2, \\ 1, & x, & x^2 \end{array}$$

are sufficient to guarantee a unique solution to the problem. Suppose a solution did not exist. Then, by Theorem (1.2.7), there exists nondegenerate algebraic curve

$$p(x, y) = \left(a_2 + b_2 y\right) x^2 + \left(a_1 + b_1 y + c_1 y^2\right) x + \left(a_0 + b_0 y + c_0\right) \quad (1.3\text{-}5)$$

which contains all the points on the above grid. Let $y = y_0$. Then $p(x, y_0)$ is a polynomial of degree 2 in the variable x which vanishes at the three points x_0, x_1 and x_2. It follows that $p(x, y_0)$ is identically zero for all values of x. In particular since $a_2 + b_2 y$ is the coefficient of x^2 we have

$$a_2 + b_2 y_0 = 0 . \tag{1.3-6}$$

In a similar way we can set $y = y_1$ in Equation (1.3-5) to obtain

$$a_2 + b_2 y_1 = 0 . \tag{1.3-7}$$

Combining Equations (1.3-6) and (1.3-7) we obtain the homogeneous system

$$\begin{aligned} a_2 + b_2 y_0 &= 0 \\ a_2 + b_2 y_1 &= 0 . \end{aligned}$$

Since $y_0 \neq y_1$ it follows that $a_2 = 0$ and $b_2 = 0$. We can use a similar technique to show the constants a_1, b_1 and c_1 are zero. In fact, if we let $y = y_0$, y_1 and y_2 and look at the coefficients for x in each case we obtain the homogeneous system

$$\begin{aligned} a_1 + y_0 b_1 + y_0^2 c_1 &= 0 \\ a_1 + y_1 b_1 + y_1^2 c_1 &= 0 \\ a_1 + y_2 b_1 + y_2^2 c_1 &= 0 . \end{aligned} \tag{1.3-8}$$

Since

$$\det\begin{pmatrix} 1 & y_0 & y_0^2 \\ 1 & y_1 & y_1^2 \\ 1 & y_2 & y_2^2 \end{pmatrix}$$

$$= \left(y_1 - y_0\right)\left(y_2 - y_1\right)\left(y_2 - y_0\right)$$

$$\neq 0 ,$$

the system (1.3-8) has the unique solution

$$a_1 = 0 ,$$

$$b_1 = 0 ,$$

$$c_1 = 0 .$$

In a similar way, one can show

$$a_0 = 0 ,$$

$$b_0 = 0 ,$$

$$c_0 = 0 .$$

Therefore, $p(x, y)$ is identically zero which contradicts the fact that $p(x, y) = 0$ was a nondegenerate algebraic curve. Therefore, by Theorem (1.2.7) we can use the monomials y^2, $y^2x, y, yx, yx^2, x, x^2$, and 1 to interpolate any function f defined in a region containing the grid in Figure (1.3.4).

In a similar way, it can be shown for the grid in Figure (1.3.5).

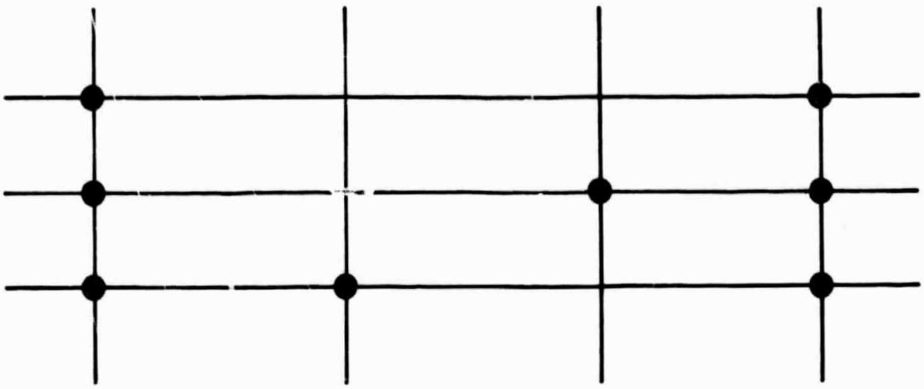

Figure (1.3.5)

A correct choice of monomials is

$$\begin{array}{lll} y^2, & y^2x^3, & \\ y, & yx^2, & yx^3, \\ 1, & x, & x^3. \end{array}$$

Or possibly

$$\begin{array}{lll} y^2, & y^2x^2, & y^2x^3, \\ yx, & yx^3, & \\ 1, & x, & x^3. \end{array}$$

1.4 GENERALIZED LINEAR INDEPENDENCE

In this section we briefly introduce the concept of generalized linear independence. In particular, we introduce the concept of homogeneous and inhomogeneous degree n independence of points and show how this idea relates

to interpolation for functions of several variables. In Section 2.4, we use this notion to determine a sufficient set of conditions under which an Aitken type interpolation scheme is valid. For notational convenience we consider points in R^{m+1}.

Definition (1.4.1): A point $\overline{x}_0 = (x_0^0, x_0^1, \ldots, x_0^m)$ in R^{m+1} is said to be a homogeneous degree n combination of the points $\overline{x}_1, \ldots, \overline{x}_p$ if there exists a collection of scalars a_j, $(j = 1, 2, \ldots, p)$ such that

$$\prod_{i=0}^{m} (x_0^i)^{n_i} = \sum_{j=1}^{p} a_j \left(\prod_{i=0}^{m} (x_j^i)^{n_i} \right) \tag{1.4-1}$$

for all sets of nonnegative integers $\{n_i : i = 0, 1, \ldots, m\}$ such that

$$\sum_{i=0}^{m} n_i = n .$$

Example (1.4.2): If $m = 2$ and $n = 2$, then a point (x_0, y_0, z_0) is a homogeneous degree 2 combination of the point (x_1, y_1, z_1) and (x_2, y_2, z_2) if and only if there exist constants a and b such that the following six equations are satisfied:

$$x_0^2 = ax_1^2 + bx_2^2$$

$$y_0^2 = ay_1^2 + by_2^2$$

$$z_0^2 = az_1^2 + bz_2^2$$

$$x_0 y_0 = ax_1 y_1 + bx_2 y_2$$

$$x_0 z_0 = ax_1 z_1 + bx_2 z_2$$

$$y_0 z_0 = ay_1 z_1 + by_2 z_2 .$$

Definition (1.4.3): A collection of points $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p$ in R^{m+1} is homogeneously degree n dependent if and only if there exist $a_1, a_2, \ldots, a_p$, not all zero, such that

$$\sum_{j=1}^{p} a_j \left(\prod_{i=0}^{m} (x_j^i)^{n_i} \right) = 0$$

for all sets of positive integers $\{n_i : i = 0, \ldots, m\}$ such that

$$\sum_{i=0}^{m} n_i = n .$$

Definition (1.4.4): A collection of points $\bar{x}_1, \ldots, \bar{x}_p$ is homogeneous degree n independent if it is not homogeneous degree n dependent.

Homogeneous degree n dependence is a natural generalization of the familiar concept of linear dependence. Many of the linear concepts such as basis and dimension have natural extensions to the more general case. For a more complete discussion the reader should consult the original work of Thacher [16].

Example (1.4.5): For the case $n = 1$ Equation (1.4-1) becomes

$$x_0^i = \sum_{j=1}^{p} a_j x_j^i ,$$

for $i = 0, 1, 2, \ldots, m$. It follows that $\bar{x}_0$ is a homogeneous degree 1 combination of the points $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p$ if and only if $\bar{x}_0$ is a linear combination of $\bar{x}_1, \bar{x}_2, \ldots \bar{x}_p$ in the usual sense.

Example (1.4.6): Consider the points

$$\bar{x}_0 = (x_0, y_0) = (1, 1)$$

$$\bar{x}_1 = (x_1, y_1) = (1, 0)$$

$$\bar{x}_2 = (x_2, y_2) = (0, 1) .$$

In this case we show $\bar{x}_0$ is not a homogeneous degree 2 combination of $\bar{x}_1$ and $\bar{x}_2$. If $\bar{x}_0$ were a degree 2 combination of $\bar{x}_1$ and $\bar{x}_2$, there would exist constants "a" and "b" such that

$$\begin{aligned} x_0^2 &= ax_1^2 + bx_2^2 \\ x_0 y_0 &= ax_1 y_1 + bx_2 y_2 \\ y_0^2 &= ay_1^2 + by_2^2 . \end{aligned} \tag{1.4-2}$$

Upon substitution, the second line of Equation (1.4-2) becomes

$$1 = a(1)(0) + b(0)(1) = 0 .$$

Therefore, the point (1, 1) cannot be a degree 2 combination of (1, 0) and (0, 1).

Many of the results for generalized linear dependence can be derived from known results in the linear case. We shall use the following well known lemma (see [6], [19]).

Lemma (1.4.7): There are $(n+m)!/m!\,n!$ monomials in $m+1$ independent variables of total degree equal to n.

Proof: A monomial in the $n+1$ variables $x^0, x^1, \ldots, x^m$ with total degree n has the form

$$\prod_{i=0}^{m} (x^i)^{n_i} ,$$

where

$$\sum_{i=0}^{m} n_i = n . \qquad (1.4\text{-}3)$$

Therefore, the number of such monomials is given by the number of nonnegative, integer valued solutions of Equation (1.4-3). It is well known that the number of nonnegative, integer valued solutions of (1.4-3) is given by $(n+m)!/n!\,m!$ (see [4], p. 36).

■

Notation: We shall use the notation

$$W_n(m) = \frac{(n+m)!}{n!\,m!} .$$

Consider the mapping

$$\phi_n : R^{m+1} \longrightarrow R^{W_n(m)}$$

defined by

$$\phi_n\left(\left(x^i\right)\right) = \left(\prod_{i=0}^{m} \left(x^i\right)^{n_i}\right),$$

where the collection $\{n_i\}$ ranges over all positive integer valued solutions of the equation

$$\sum_{i=0}^{m} n_i = n .$$

Example (1.4.8): For $n = 2$ and $m = 2$ we have

$$\phi_2 : R^3 \longrightarrow R^6$$

defined by

$$\phi_2 (x, y, z) = \left(x^2, y^2, z^2, xy, yz, xz\right) .$$

The following theorem relates homogeneous degree n independence in R^{m+1} to linear independence in $R^{W_n(m)}$.

Theorem (1.4.9): Let $\overline{x}_1, \ldots, \overline{x}_p$ be a collection of points in R^{m+1}. Then $\overline{x}_1, \ldots, \overline{x}_p$ is homogeneous degree n independent if and only if the collection $\phi_n(\overline{x}_1), \ldots, \phi_n(\overline{x}_p)$ is linearly independent in $R^{W_n(m)}$.

Proof: Suppose $\phi_n \overline{x}_1, \ldots, \phi_n \overline{x}_p$ is linearly independent in $R^{W_n(m)}$. Then, there are constants $a_1, \ldots, a_p$, not all zero, such that

$$\sum_{j=1}^{p} a_j \phi_n\left(\overline{x}_j\right) = 0 .$$

Looking at the components of each $\phi_n\left(\overline{x}_j\right)$ we have

$$\sum_{j=1}^{p} a_j \left(\prod_{i=0}^{m} \left(x_j^{\,i}\right)^{n_i}\right) = 0$$

for all nonnegative integer solutions $\{n_i\}$ of the equation

$$\sum_{i=0}^{m} n_i = n .$$

It follows immediately from the definition that the points $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p$ are homogeneously degree n independent in R^{m+1}. The converse follows in the same way.

■

In order to relate the above material to multivariate interpolation we need the following definition.

Definition (1.4.10): A collection of points $(\bar{x}_i)$ in R^m is inhomogeneously degree n dependent (independent) if the collection

$$\bar{x}_i' = \left(1, x_i^1, x_i^2, \cdots, x_i^m\right)$$

is homogeneously degree n dependent (independent) in R^{m+1}.

We can now present the main result of this section:

Theorem (1.4.11): Let f be a real valued function defined in a region G of R^m and let $\{\bar{x}_i\}$ be a collection of $W_n(m)$ points in G. Then there exists a polynomial ϕ of total degree n such that

$$\phi(\bar{x}_i) = f(\bar{x}_i)$$

for $i = 1, 2, \ldots, W_n(m)$, if and only if the collection is inhomogeneously degree n independent.

Proof: Let $\{\bar{x}_i\}$ be a collection of $W_n(m)$ inhomogeneously degree n independent points in R^m. By Theorem (1.4.9), the vectors $\left\{\phi_n\left(\bar{x}_i'\right)\right\}$ are linearly independent in $R^{W_n(m)}$. But since the vectors $\left\{\phi_n\left(\bar{x}_i'\right)\right\}$ are the rows of the determinant in Equation (1.2-4) when the basis functions are taken as all monomials in m variables of total degree less than or equal to n, the existence of the polynomial ϕ follows immediately from Theorem (1.2.1). The converse follows in a similar manner.

■

Corollary (1.4.11): A collection of $W_n(m)$ points $\{\bar{x}_i\}$ in R^m is inhomogeneously degree n independent if and only if there exists a nondegenerate algebraic curve $\phi(\bar{x}) = 0$ with total degree n, such that

$$\phi(\bar{x}_i) = 0$$

for $i = 1, 2, \ldots, W_n(m)$.

For some further geometric results for degree n independence the reader should consult the paper by Whaples [20].

1.5 INTERPOLATION IN VECTOR SPACES

Let $\phi_0, \phi_1, \ldots, \phi_n$ be a collection of functions defined in a region G of R^m and let V be the vector space of functions generated by this collection. If $\bar{y}$ is a point in R^m, the evaluation function defined by

$$E_{\bar{y}}(\phi) = \phi(\bar{y}) \qquad (1.5\text{-}1)$$

is a linear operator from the vector space V to the real numbers. In fact if ϕ and ψ are members of V and if c_1 and c_2 are real numbers we have

$$\begin{aligned} E_{\bar{y}}\left(c_1\phi + c_2\psi\right) &= \left(c_1\phi + c_2\psi\right)(\bar{y}) \\ &= c_1\phi(\bar{y}) + c_2\psi(\bar{y}) \\ &= c_1 E_{\bar{y}}(\phi) + c_2 E_{\bar{y}}(\psi) . \end{aligned}$$

It is not necessary to restrict our attention to linear functionals defined by Equation (1.5-1). In fact any collection of m + 1 linearly independent functionals on V is sufficient to guarantee a unique solution to the interpolation problem. For example, if the functions $\{\phi_i\}$ and f are sufficiently smooth, we might consider a linear functional defined by

$$L(\phi) = \frac{\partial^2(\phi)}{\partial x_i\, \partial x_j}\left(\bar{x}_0\right)$$

or possibly one in the form

$$L(\phi) = \int_G \phi\, d\bar{x} .$$

In fact the above methods can be generalized considerably. For arbitrary vector spaces Theorem (1.2.1) takes the following form:

Theorem (1.5.1): Let V be an $m+1$ dimensional vector spare with basis $\{v_0, v_1, \ldots, v_m\}$ and let $L_0, L_1, \ldots, L_m$ be $m+1$ linear functionals defined in V. The following conditions are equivalent:

1) The collection $\{L_j\}$ is linearly independent.

2) For all collections of real numbers $r_0, r_1, \ldots, r_m$, the condition

$$L_j(v) = r_j$$

for $j = 0, 1, \ldots, m$ defines a unique vector in V.

3) The determinant $\det\left[L_j(v_i)\right]$ is not zero.

Proof: (1) $\Longrightarrow$ (3). Suppose $\det\left[L_j(v_i)\right] \neq 0$. Then the homogeneous system

$$\sum_{j=0}^{m} a_j L_j(v_1) = 0$$

$$\vdots$$

$$\sum_{j=0}^{m} a_j L_j(v_m) = 0$$

has a nontrivial solution. Since

$$\sum_{j=0}^{m} a_j L_j(v_i) = 0$$

for $i = 0, 1, \ldots, m$, it follows that

$$\sum_{j=0}^{m} a_j L_j(v) = 0$$

for all v in V. However, this means the collection $\{L_j\}$ is linearly dependent which contradicts (1). Therefore $\det\left[L_j\,(v_i)\right] = 0$.

(3) $\Longrightarrow$ (2). A vector v in V has the form

$$\sum_{i=0}^{m} a_i\, v_i\,.$$

Consider the system

$$\begin{aligned} \sum_{i=0}^{m} a_i\, L_1\,(v_i) &= r_1 \\ &\vdots \\ \sum_{i=0}^{m} a_i\, L_m\,(v_i) &= r_m \end{aligned} \qquad (1.5\text{-}2)$$

Since

$$\det\left[L_j\,(v_i)\right] \neq 0\,,$$

the system (1.5-2) has a unique solution for the coefficients $a_0, a_1, \ldots, a_m$. The unique vector

$$v = \sum_{i=0}^{m} a_i\, v_i$$

satisfies the property

$$L_j(v) = r_j$$

for $j = 0, 1, \ldots, m$.

(2) $\Longrightarrow$ (1). Suppose the collection $\{L_i\}$ is not linearly independent. Without loss of generality we assume

$$L_0 = \sum_{j=1}^{m} c_j L_j . \tag{1.5-3}$$

For a collection of real numbers $r_0, \ldots, r_m$ we have a unique vector v in V such that

$$L_j(v) = r_j .$$

Furthermore, by Equation (1.5-3)

$$L_0(v) = \sum_{j=1}^{m} c_j r_j . \tag{1.5-4}$$

Choose a new set of constants $s_0, \ldots, s_m$ such that

$$s_0 \neq r_0 .$$

and

$$s_j = r_j$$

for $1 \leq j \leq m$. Because of Equation (1.5-4) we cannot find a vector v such that

$$L_j(v) = s_j$$

for $0 \leq j \leq m$. Therefore, the collection $\{L_j\}$, is linearly independent. ■

CHAPTER II

INTERPOLATION FORMULAE FOR SEVERAL VARIABLES

2.1 THE FUNDAMENTAL FORMULA

Almost all interpolation formulas can be derived from the following elementary theorem on finite dimensional vector spaces.

Theorem (2.1.1): (Fundamental Theorem of Interpolation Theory) Let V be an $n+1$ dimensional vector space with basis $v_0, v_1, \ldots, v_n$ and let $L_0, \ldots, L_n$ be a basis for the dual space V^*. Then any v in V can be expressed in the form

$$v = -\frac{\det \begin{bmatrix} L_0(v_0) & \cdots & L_n(v_0) & v_0 \\ \vdots & & \vdots & \vdots \\ L_0(v_n) & \cdots & L_n(v_n) & v_n \\ L_0(v) & \cdots & L_n(v) & 0 \end{bmatrix}}{\det\left[L_i(v_j)\right]} . \qquad (2.1\text{-}1)$$

Proof: The expression on the right is a linear combination of $v_0, \ldots, v_n$ and hence is a member of V. Let us denote this expression by v'. If we apply the linear functional L_i to v' we obtain

$$L_i(v') = -\frac{\det \begin{bmatrix} L_0(v_0) & \cdots & L_n(v_0) & L_i(v_0) \\ \vdots & & \vdots & \vdots \\ L_0(v_n) & \cdots & L_n(v_n) & L_i(v_n) \\ L_0(v) & \cdots & L_n(v) & 0 \end{bmatrix}}{\det\left[L_i(v_j)\right]} .$$

In the determinant in the numerator we subtract the i^{th} column from the last column and obtain

$$L_i(v') = -\frac{\det\begin{bmatrix} L_0(v_0) & \cdots & L_n(v_0) & 0 \\ \vdots & & \vdots & \vdots \\ L_0(v_n) & \cdots & L_n(v_n) & 0 \\ L_0(v) & \cdots & L_n(v) & -L_i(v) \end{bmatrix}}{\det\left[L_i(v_j)\right]}.$$

Expanding the numerator by the last column we obtain

$$L_i(v') = \frac{-(-L_i(v))\det\left[L_i(v_j)\right]}{\det\left[L_i(v_j)\right]}$$

$$= L_i(v),$$

for $i = 0, 1, 2, \ldots, n$. By Theorem (1.5.1-2) we have

$$v = v'.$$

■

By interchanging the roles of V and V^* we obtain

<u>Theorem (2.1.2)</u>: Let V^* be the dual space of an $n+1$ dimensional vector space V. Let $L_0, \ldots, L_n$ be a basis in V^* and $v_0, \ldots, v_n$ be a basis in V. Then any L in V^* can be written

$$L = -\frac{\det\begin{bmatrix} L_0(v_0) & \cdots & L_0(v_n) & L_0 \\ \vdots & & \vdots & \vdots \\ L_n(v_0) & \cdots & L_n(v_n) & L_n \\ L(v_0) & \cdots & L(v_n) & 0 \end{bmatrix}}{\det\left[L_i(v_j)\right]}.$$

In the determinant in the numerator we subtract the i^{th} column from the last column and obtain

$$L_i(v') = -\frac{\det\begin{bmatrix} L_0(v_0) & \cdots & L_n(v_0) & 0 \\ \vdots & & \vdots & \vdots \\ L_0(v_n) & \cdots & L_n(v_n) & 0 \\ L_0(v) & \cdots & L_n(v) & -L_i(v) \end{bmatrix}}{\det\left[L_i(v_j)\right]}.$$

Expanding the numerator by the last column we obtain

$$L_i(v') = \frac{-(-L_i(v))\det\left[L_i(v_j)\right]}{\det\left[L_i(v_j)\right]}$$

$$= L_i(v),$$

for $i = 0, 1, 2, \ldots, n$. By Theorem (1.5.1-2) we have

$$v = v'.$$

■

By interchanging the roles of V and V^* we obtain

<u>Theorem (2.1.2)</u>: Let V^* be the dual space of an $n + 1$ dimensional vector space V. Let $L_0, \ldots, L_n$ be a basis in V^* and $v_0, \ldots, v_n$ be a basis in V. Then any L in V^* can be written

$$L = -\frac{\det\begin{bmatrix} L_0(v_0) & \cdots & L_0(v_n) & L_0 \\ \vdots & & \vdots & \vdots \\ L_n(v_0) & \cdots & L_n(v_n) & L_n \\ L(v_0) & \cdots & L(v_n) & 0 \end{bmatrix}}{\det\left[L_i(v_j)\right]}.$$

In other words, the Lagrange coefficients $\ell_i(x)$ are the unique dual basis in $V^{**} = V$ associated with the basis $E_0, E_1, \ldots, E_n$ in V^*. It is actually possible to obtain a formula for a general vector space which is similar to Equation (2.2-1). If we expand the determinant in the numerators of Equation (2.1-2) by the bottom row and then rearrange the columns we obtain

$$v = \sum_{i=0}^{n} L_i(v)\psi_i \tag{2.2-3}$$

where

$$\psi_i = \frac{\det\begin{bmatrix} L_0(v_0) & \cdots & L_{i-1}(v_0) & v_0 & L_{i+1}(v_0) & \cdots & L_n(v_0) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ L_0(v_n) & \cdots & L_{i-1}(v_n) & v_n & L_{i+1}(v_n) & \cdots & L_n(v_n) \end{bmatrix}}{\det\left[L_i(v_j)\right]} . \tag{2.2-4}$$

Formula (2.2-3) is known as the generalized Lagrange interpolation formula and the ψ_i are the generalized Lagrange interpolation coefficients. One can easily check that

$$L_i(\psi_j) = \delta_{ij} .$$

For polynomial interpolation in several variables the Lagrange formula takes the following form:

Theorem (2.2.1): Let f be a function defined in a region G of R^n and let $\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_m$ be $m+1$ points in G. Let $\phi_0(\bar{x}), \phi_1(\bar{x}), \ldots, \phi_m(\bar{x})$ be $m+1$ monomials in n variables such that

$$\det\left[\phi_i(x_j)\right] \neq 0 .$$

Then, the unique interpolation polynomial is given by

$$p(\overline{x}) = \sum_{j=0}^{m} f(\overline{x}_j)\, \ell_j(\overline{x}) \,, \tag{2.2-5}$$

where

$$\ell_j(\overline{x}) = \frac{\det \begin{bmatrix} \phi_0(\overline{x}_0) & \cdots & \phi_0(\overline{x}_{j-1}) & \phi_0(\overline{x}) & \phi_0(\overline{x}_{j+1}) & \cdots & \phi_0(\overline{x}_n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \phi_n(\overline{x}_0) & \cdots & \phi_n(\overline{x}_{j-1}) & \phi_n(\overline{x}) & \phi_n(\overline{x}_{j+1}) & \cdots & \phi_n(\overline{x}_n) \end{bmatrix}}{\det\left[\phi_i(\overline{x}_j)\right]} \,.$$

Proof: The proof is an application of Theorem (1.5.1) and the preceding discussion.

■

In the case of single variable interpolation, we can use the properties of Vandermonde determinants to reduce the Lagrange coefficients to the form found in Equation (2.2-1). Except for special cases, such a reduction for the several variables case is not possible.

Lagrange Interpolation on a Two Dimensional Grid

One case, where a simplification of the Lagrange coefficients is possible, occurs when the points are located on a rectangular grid as shown in Figure (2.2.1).

Let f be a function defined in a region G containing the above grid. By Theorem (1.3.2), we can obtain a unique solution to the interpolation problem by using the monomials

$$\begin{matrix} 1\,, & x\,, & \cdots\,, & x^n \\ y\,, & yx\,, & \cdots\,, & yx^n \\ \vdots & \vdots & & \vdots \\ y^m\,, & y^m x\,, & \cdots\,, & y^m x^n \,. \end{matrix} \tag{2.2-7}$$

Then, the unique interpolation polynomial is given by

$$p(\bar{x}) = \sum_{j=0}^{m} f(\bar{x}_j)\, \ell_j(\bar{x}) \,, \tag{2.2-5}$$

where

$$\ell_j(\bar{x}) = \frac{\det \begin{bmatrix} \phi_0(\bar{x}_0) & \cdots & \phi_0(\bar{x}_{j-1}) & \phi_0(\bar{x}) & \phi_0(\bar{x}_{j+1}) & \cdots & \phi_0(\bar{x}_n) \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \phi_n(\bar{x}_0) & \cdots & \phi_n(\bar{x}_{j-1}) & \phi_n(\bar{x}) & \phi_n(\bar{x}_{j+1}) & \cdots & \phi_n(\bar{x}_n) \end{bmatrix}}{\det\left[\phi_i(\bar{x}_j)\right]} \,.$$

Proof: The proof is an application of Theorem (1.5.1) and the preceding discussion.

■

In the case of single variable interpolation, we can use the properties of Vandermonde determinants to reduce the Lagrange coefficients to the form found in Equation (2.2-1). Except for special cases, such a reduction for the several variables case is not possible.

Lagrange Interpolation on a Two Dimensional Grid

One case, where a simplification of the Lagrange coefficients is possible, occurs when the points are located on a rectangular grid as shown in Figure (2.2.1).

Let f be a function defined in a region G containing the above grid. By Theorem (1.3.2), we can obtain a unique solution to the interpolation problem by using the monomials

$$\begin{matrix} 1\,, & x\,, & \cdots\,, & x^n \\ y\,, & yx\,, & \cdots\,, & yx^n \\ \vdots & \vdots & & \vdots \\ y^m\,, & y^m x\,, & \cdots\,, & y^m x^n\,. \end{matrix} \tag{2.2-7}$$

The polynomial $p(x, y)$ is a linear combination of the monomials in (2.2-7) and it satisfies the property

$$p(x_i, y_j) = f(x_i, y_j) .$$

Furthermore, the coefficients $\ell_i(x)\ell_j(y)$ satisfy the fundamental property

$$\ell_i(x_i)\ell_j(y_j) = 1$$

$$\ell_i(x_k)\ell_j(y_\ell) = 0 \qquad k \neq i, \ \ell \neq j .$$

Alternately, we may derive Equation (2.2-8) from Equation (2.2-6). Although this approach is more complicated we include it to illustrate the techniques needed to derive results from the general formula (2.2-6). We shall use the following lemma.

Lemma (2.2.3): Let (x_i, y_j) $(i = 0, 1, \ldots, n)$, $(j = 0, 1, \ldots, m)$ be the $(m+1)(n+1)$ points on the grid in Figure (2.2.1). Then for each fixed set of indices (i^1, j^1) we have

$$\det \begin{bmatrix} 1 & x_0 & y_0 & x_0^2 & \cdots & x_0^n y_0^m \\ 1 & x_1 & y_0 & x_1^2 & \cdots & x_1^n y_0^m \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & y_m & x_n^2 & \cdots & x_n^n y_m^m \end{bmatrix}$$

$$= \prod_{\substack{i=0 \\ i \neq i1}}^{n} \left(x_{i^1} - x_i\right) \prod_{\substack{j=0 \\ j \neq j1}}^{m} \left(y_{j^1} - y_j\right) \mathrm{Cof}\left(x_{i^1}^m \, y_{j^1}^n\right) .$$

Proof: Consider the matrix

$$P = \begin{vmatrix} 1 & x_0 & y_0 & x_0^2 & \cdots & x_0^n y_0^m \\ 1 & x_1 & y_0 & x_1^2 & \cdots & x_1^n y_0^m \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & y_m & x_n^2 & \cdots & x_n^n y_m^m \end{vmatrix} . \qquad (2.2\text{-}9)$$

Replace the pair $(x_{i'}, y_{i'})$ in P by the variable points (x, y) and let $P(x, y)$ denote the polynomial obtained by taking $\det[P]$ with this substitution. Let $x = x_0$. It follows that $P(x_0, y)$ is a polynomial of degree m in the variable y which vanishes at the $m + 1$ points $y_0, y_1, \ldots, y_m$ and therefore for all y. Therefore, by Lemma (1.3.1), there exists a polynomial $Q_0(x, y)$ such that

$$P(x, y) = (x - x_0) Q_0(x, y) .$$

Repeating the above procedure for each $i \neq i^1$ and each $j \neq j^1$ we obtain

$$P(x, y) = \prod_{i \neq i'} (x - x_i) \prod_{j \neq j'} (y - y_j) \, C(x, y) .$$

To determine $C(x, y)$ we note that

$$\prod_{i \neq i'} (x - x_i) \prod_{j \neq j'} (y - y_j)$$

contains a term $x^n y^m$ and therefore $C(x, y)$ must be the coefficient of $x^n y^m$ in $P(x, y)$. Expanding the determinant of the matrix which defines $P(x, y)$ by the

last column we see that

$$C(x, y) = \text{Cof}\left(x_i^n, y_j^m\right) .$$

Therefore

$$P(x, y) = \prod_{i \neq i'} (x - x_i) \prod_{j \neq j'} (y - y_j) \, \text{Cof}\left(x_{i1}^n \, y_{j1}^m\right) .$$

Letting $x = x_{i'}$ and $y = y_{j'}$ we obtain the result of the lemma.

Theorem (2.2.4): The Lagrange coefficients for interpolation on a two dimensional grid are given by

$$\ell_{i^1 j^1}(x, y) = \frac{\prod_{i \neq i'} (x - x_i) \prod_{j \neq j'} (y - y_j)}{\prod_{i \neq i'} \left(x_{i1} - x_i\right) \prod_{j \neq j'} \left(y_{j1} - y_j\right)} .$$

Proof: Evaluating the determinants for the Lagrange coefficients in Equation (2.2-6) by the previous lemma we have

$$\ell_{i' j'}(x, y) = \frac{\prod_{i \neq i'} (x - x_i) \prod_{j \neq j'} (y - y_j) \, \text{Cof}\left(x^n \, y^m\right)}{\prod_{i \neq i'} \left(x_{i1} - x_i\right) \prod_{j \neq j'} \left(y_{j1} - y_j\right) \text{Cof}\left(x_{i1}^n \, y_{j1}^m\right)} .$$

However, since

$$\mathrm{Cof}\left(x^n y^m\right) = \mathrm{Cof}\left(x_i^n, y_j^m\right) \neq 0,$$

we obtain the result. ∎

2.3 DIVIDED DIFFERENCES FOR SEVERAL VARIABLES

In this section we present a generalization to functions of several variables of Newton's fundamental interpolation formula in terms of divided differences. The early definitions for divided differences of functions of several variables (see [8]) where based on an iteration of the difference operations for each variable taken separately. By there very nature, these processes place the base points at the corners of an n-dimensional rectangular grid. A different definition is given by Whittaker and Robinson [21] where the points are located on a two dimensional triangular grid. Salzer (see [10]) presents two more general schemes, however neither scheme gives a polynomial of lowest possible degree and in both cases we obtain polynomials with more coefficients than base points.

The generalization of divided differences that we present was first explicitly given by Salzer [11], however a more general scheme which can easily be applied to the several variable case had been given earlier by Curry [1] and more recently by Davis [3]. We present a variation of Curry's method and apply it to polynomial interpolation in several variables to obtain some of Salzer's results. The generalization presented here is applicable to arbitrarily located points as long as the determinants appearing in the denominators are not zero.

To motivate the several dimensional approach we briefly list some of the main properties of divided differences and Newton's fundamental formula.

Let f be a function defined on an interval [a, b] which contains the points $x_0, x_1, \ldots, x_n$. Since the polynomials

$$V_0 = 1$$

$$V_1 = \left(x - x_0\right)$$

$$V_n = \prod_{i=0}^{n-1} \left(x - x_i\right)$$

are linearly independent, they form a basis for the vector space V spanned by the monomials $1, x, x^2, \ldots, x^n$. Let $E_i(\phi)$ be the evaluation functional on V defined by

$$E_i(\phi) = \phi(x_i) .$$

Since

$$\det\left[E_i\left(V_j\right)\right] = \det\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & \left(x_1 - x_0\right) & \cdots & \left(x_n - x_0\right) \\ 0 & 0 & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \prod_{i=0}^{n-1}\left(x_n - x_i\right) \end{bmatrix}$$

$$\neq 0 ,$$

it follows from Theorem (1.5.1) that there exists a polynomial

$$p(x) = \sum_{j=0}^{n} a_j V_j(x) \tag{2.3-1}$$

such that

$$p\left(x_i\right) = f\left(x_i\right)$$

for $i = 0, 1, 2, \ldots, n$. In Equation (2.3-1) the coefficients a_j are called the divided differences of order j and are usually denoted by $[0, 1, \ldots, j]$, $[x_0, \ldots, x_j]$ or $[f(x_0), \ldots, f(x_j)]$. The polynomials $\{V_j(x)\}$ are known as the Newton Polynomials.

Three important properties of these differences are summarized in the following theorem.

Theorem (2.3.1):

1) The divided difference [0, 1, . . . , j] can be considered as a linear functional on

$$V = sp\{1, x, x^2, \cdots, x^n\}$$

which is a linear combination of the evaluation functionals $E_1, E_2, \ldots, E_j$. It is given explicitly by

$$[0, 1, \cdots, j](\phi) = \frac{\det\begin{bmatrix} 1 & x_0 & \cdots & x_0^{j-1} & E_0(\phi) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & x_j & \cdots & x_j^{j-1} & E_j(\phi) \end{bmatrix}}{\det\begin{bmatrix} 1 & x_0 & \cdots & x_0^j \\ \vdots & \vdots & & \vdots \\ 1 & x_j & \cdots & x_j^j \end{bmatrix}} .$$

2) The divided difference [0, 1, 2, . . . , j] satisfies the following:

$$\begin{aligned}[0, 1, \cdots, j]\left(x^k\right) &= 0 \qquad k < j \\ &= 1 \qquad k = j .\end{aligned}$$

3) The divided difference can be computed by the following recursion formula

$$[0, 1, \cdots, j] = \frac{[1, 2, \cdots, j] - [0, 1, \cdots, j-1]}{x_j - x_0} .$$

Proof: Statement (1) is proved in Householder ([7], p. 203) and is also a special case of Theorem (2.3.3). Statement (2) follows directly from statement (1). Statement (3) is discussed in almost any book on numerical analysis. ■

In a similar way the Newton polynomials possess a number of important properties.

Theorem (2.3.2):

1) $E_j\,(V_k) \;=\; 0 \qquad\qquad \text{for} \qquad\qquad j < k\,.$

2) V_j is a linear combination of $1, x, x^2, \ldots, x^j$.

3) $$L_j\left(V_j\right) \;=\; \prod_{i=0}^{j-1}\left(x_j - x_i\right) \;=\; \frac{\det\begin{bmatrix} 1 & x_0 & \cdots & x_0^{\,j} \\ \vdots & \vdots & & \vdots \\ 1 & x_j & \cdots & x_j^{\,j} \end{bmatrix}}{\det\begin{bmatrix} 1 & x_0 & \cdots & x_0^{\,j-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{j-1} & \cdots & x_{j-1}^{\,j-1} \end{bmatrix}}\,.$$

Proof: All three properties follow directly from the definition and the properties of Vandermonde determinants. ■

With this motivation we have the following general theorem in finite dimensional vector spaces.

Theorem (2.3.3): Let V be a finite dimensional vector space with basis V_0, $V_1, \ldots, V_n$. Let $L_0, L_1, \ldots, L_n$ be a collection of linear functionals on V such that the subcollection $L_0, L_1, \ldots, L_k$ is linearly independent when considered as linear functionals over

$$V^k \;=\; \mathrm{sp}\left\{V_0,\, V_1,\, \cdots,\, V_k\right\}\,.$$

Then, there exists a unique basis $N_0, N_1, \ldots, N_n$ for V^* satisfying the properties

a-1) N_k is a linear combination of $L_0, \ldots, L_k$

a-2) $N_k\left(V_j\right) = 0 \qquad\qquad \text{for} \qquad\qquad j < k$

a-3) $N_k\left(V_j\right) = 1.$

Similarly, there exists a unique basis $W_0, \ldots, W_n$ of V which satisfies the following properties

b-1) W_k is a linear combination of $L_0, \ldots, L_k$

b-2) $L_j\left(W_k\right) = 0$ for all $j < k$

b-3) $$L_k\left(W_k\right) = \frac{\det\begin{bmatrix} L_0\left(V_0\right) & \cdots & L_k\left(V_0\right) \\ \vdots & & \vdots \\ L_0\left(V_k\right) & \cdots & L_k\left(V_k\right)\end{bmatrix}}{\det\begin{bmatrix} L_0\left(V_0\right) & \cdots & L_{k-1}\left(V_0\right) \\ \vdots & & \\ L_0\left(V_{k-1}\right) & \cdots & L_{k-1}\left(V_{k-1}\right)\end{bmatrix}} . \qquad (2.3\text{-}2)$$

The functionals N_k are known as the <u>generalized divided differences</u> for V and the vectors W_k in the <u>generalized Newton Polynomials.</u>

<u>Proof:</u> We shall just determine the generalized divided differences N_k. Consider the subspace

$$V^k = sp\left\{V_0, V_1, \cdots, V_k\right\} .$$

By our hypothesis, the functional $L_0, \ldots, L_k$ are linearly independent when considered as functionals over V^k. By Theorem (2.1.2), there exists a unique vector N_k in V_k^* satisfying the property

$$N_k\left(V_j\right) = 0 \quad \text{for} \quad j < k$$

$$N_k\left(V_k\right) = 1 .$$

In fact N_k is given by

$$N_k = - \frac{\det\begin{bmatrix} L_0(V_0) & \cdots & L_0(V_k) & L_0 \\ \vdots & & \vdots & \vdots \\ L_k(V_0) & \cdots & K_k(V_k) & L_n \\ 0 & \cdots 0 & 1 & 0 \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_0(V_k) \\ \vdots & & \vdots \\ L_k(V_0) & \cdots & L_k(V_k) \end{bmatrix}}.$$

Expanding the numerator of N_k by the bottom row we obtain

$$N_k = \frac{\det\begin{bmatrix} L_0(V_0) & \cdots & L_0(V_{k-1}) & L_0 \\ \vdots & & \vdots & \vdots \\ L_k(V_0) & \cdots & L_k(V_{k-1}) & L_k \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_0(V_k) \\ \vdots & & \vdots \\ L_k(V_0) & \cdots & L_k(V_k) \end{bmatrix}}. \qquad (2.3\text{-}3)$$

To obtain the Newton polynomials we apply the exact same process. Since the functionals $L_0, \ldots, L_k$ are a basis for $(V^k)^*$ we can apply Theorem (2.1.1) to obtain the existence of a vector W_k in V such that conditions (b-2) and (b-3) are satisfied. In fact

$$W_k = - \frac{\det\begin{bmatrix} L_0(V_0) & \cdots & L_k(V_0) & V_0 \\ \vdots & & \vdots & \vdots \\ L_0(V_k) & \cdots & L_k(V_k) & V_k \\ 0 \; 0 & \cdots 0 & L_k(W_k) & 0 \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_k(V_0) \\ \vdots & & \vdots \\ L_0(V_k) & \cdots & L_k(V_k) \end{bmatrix}}.$$

Expanding the numerator by the bottom row we obtain

$$W_k = L_k(W_k)\,\frac{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{k-1}(V_0) & V_0 \\ \vdots & & \vdots & \vdots \\ L_0(V_k) & \cdots & L_{k-1}(V_k) & V_k \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_k(V_0) \\ \vdots & & \vdots \\ L_0(V_k) & \cdots & L_k(V_k) \end{bmatrix}} .$$

Using Equation (2.3-2) we obtain

$$W_k = \frac{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{k-1}(V_0) & V_0 \\ \vdots & & \vdots & \vdots \\ L_0(V_k) & \cdots & L_{k-1}(V_k) & V_k \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{k-1}(V_0) \\ \vdots & & \vdots \\ L_0(V_{k-1}) & \cdots & L_{k-1}(V_{k-1}) \end{bmatrix}} . \qquad (2.3\text{-}4)$$

Conditions (a-1), (a-2), (a-3) and (b-1), (b-2) and (b-3) follow from the expressions for N_k and W_k.

In order to show $\{N_k\}$ and $\{W_k\}$ are bases for V^* and V we must show they are linearly independent. Suppose the collection $\{N_k\}$ is not linearly independent. Then there exist constants $\{a_i\}$ not all zero such that

$$\sum_{i=0}^{n} a_i N_i = 0 .$$

Let i_0 be the maximum index such that $a_{i_0} \neq 0$. Then, if $c_i = a_i / a_{i_0}$, we can write

$$N_{i_0} = -\sum_{i=0}^{i_0 - 1} c_i N_i .$$

But this implies N_{i_0} is a linear combination of $L_1, L_2, \ldots, L_{i_0-1}$. However, using Equation (2.3-3) this will imply L_{i_0} is a linear combination of $L_1, L_2, \ldots, L_{i_0-1}$ which is impossible. Therefore $\{N_k\}$ is linearly independent. In a similar way we can show $\{W_k\}$ is linearly independent.

■

In the next theorem we show $\{N_k\}$ is the unique dual basis associated with $\{W_k\}$.

Theorem (2.3.4): The divided differences $\{N_k\}$ and the Newton polynomials $\{W_k\}$ satisfy the following relationship:

$$N_k(W_j) = \delta_{jk}.$$

Proof: Let W_{j_0} be a fixed Newton polynimial. We divide the proof into three cases:

1) Case I. $(k > j_0)$. By Theorem (2.3.3) we have

$$N_k(V_j) = 0$$

for $j < k$. Since W_{j_0} is a linear combination of $V_1, V_2, \ldots, V_{j_0}$ and since $j_0 < k$ we have

$$N_k\left(V_{j_0}\right) = 0.$$

2) Case II. $(k = j_0)$. Using properties (a-2) and (a-3) in Theorem (2.3.3), formula (2.3-4) gives

$$N_{j_0}\left(W_{j_0}\right) = \frac{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{k-1}(V_0) & 0 \\ \vdots & & \vdots & \vdots \\ L_0(V_k) & \cdots & L_{k-1}(V_k) & 1 \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{k-1}(V_0) \\ \vdots & & \vdots \\ L_0(V_{k-1}) & \cdots & L_{k-1}(V_{k-1}) \end{bmatrix}}$$

$$= 1.$$

3) <u>Case III. $(k < j_0)$</u>. By Theorem (2.3.3) we can write N_k in the form

$$N_k = \sum_{i=0}^{k} c_i L_i .$$

Therefore, using formula (2.3-4) we obtain

$$N_k\left(W_{j_0}\right) = \frac{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{j_0-1}(V_0) & \sum_{i=0}^{k} c_i L_i(V_0) \\ \vdots & & \vdots & \vdots \\ L_0\left(V_{j_0}\right) & \cdots & L_{j_0-1}\left(V_{j_0}\right) & \sum_{i=0}^{k} c_i L_i\left(V_{j_0}\right) \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{j_0-1}(V_0) \\ \vdots & & \vdots \\ L_0\left(V_{j_0-1}\right) & \cdots & L_{j_0-1}\left(V_{j_0-1}\right) \end{bmatrix}}$$

$$= \frac{\sum_{i=0}^{k} c_i \det\begin{bmatrix} L_0(V_0) & \cdots & L_{j_0-1}(V_0) & L_i(V_0) \\ \vdots & & \vdots & \vdots \\ L_0\left(V_{j_0}\right) & \cdots & L_{j_0-1}\left(V_{j_0}\right) & L_i\left(V_{j_0}\right) \end{bmatrix}}{\det\begin{bmatrix} L_0(V_0) & \cdots & L_{j_0-1}(V_0) \\ \vdots & & \vdots \\ L_0\left(V_{j_0-1}\right) & \cdots & L_{j_0-1}\left(V_{j_0-1}\right) \end{bmatrix}} .$$

For each determinant in the numerator above we have two columns alike; therefore

$$N_k\left(W_{j_0}\right) = 0 .$$

■

We are now able to obtain the generalized Newton representation formula for both vectors in V and functionals in V^*.

Theorem (2.3.5): Let V, $\{V_i\}$ and $\{L_i\}$ satisfy the same properties as in Theorem (2.3.3). Then for all u in V we have

$$u = \sum_{i=0}^{n} N_k(u)\, W_k . \tag{2.3-5}$$

For all L in V^* we have

$$L = \sum_{k=0}^{n} L\left(W_k\right) N_k . \tag{2.3-6}$$

Proof: We prove (2.3-5). Formula (2.3-6) is proved in the same way. By Theorems (2.1.1) and (2.3.3) we have

$$u = - \frac{\det \begin{bmatrix} 0 & 1 & \cdots & 0 & W_0 \\ 0 & 1 & \cdots & 0 & W_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & & \cdots & 1 & W_n \\ N_0(u) & & \cdots & N_n(u) & 0 \end{bmatrix}}{\det \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}} .$$

Evaluating the determinant we have the result. ∎

Before we derive a recursion formula analogous to part (3) of Theorem (2.3.1), we make a number of observations. From Equation (2.3-4) we observe that W_k depends on the vector space basis $V_0, \ldots, V_k$ and the functionals $L_0, \ldots, L_k$ and it is invariant under any permutation of these functionals. In the proof below

$$W_k\left(a_0, a_1, \cdots, a_{k-1}\right)$$

to indicate the Newton polynomial depends on the basis $V_0, V_1, \ldots, V_k$ and the linear functionals $L_{a_1}, \ldots, L_{a_k}$. Similarly, from Equation (2.3-3) we can see N_k depends on the basis vectors $V_0, V_1, \ldots, V_k$ and the functionals $L_0, L_1, \ldots, L_k$. We shall use the notation

$$N_k\left(a_0, a_1, \cdots, a_k\right)$$

to indicate dependence on $V_0, V_1, \ldots, V_k$ and $L_{a_0}, \ldots, L_{a_k}$.

We can now prove the promised recursion formula. Throughout this discussion we have an ordered basis $\{V_k\}$ for V which shall remain fixed. Assume $k \geq 2$. We shall derive a formula for $N_k(0, 1, \ldots, k)$ in terms of $N_{k-1}(0, 1, \ldots, k-1)$ and $N_{k-1}(1, \ldots, k)$. Let

$$L_1, L_2, \cdots, L_{k-1}, L_0, L_k, L_{k+1} \cdots, L_n$$

and

$$L_1, L_2, \cdots, L_{k-1}, L_k, L_0, L_{k+1}, \cdots, L_n$$

be two orderings of the linear functionals $\{L_i\}$ and let u belong to V. Using Theorem (2.3.5) and the first ordering we can write

$$\begin{aligned} u = N_0(1) V_0 &+ N_1(1, 2) W_1(1) + \cdots \\ &+ N_{k-2}(1, \cdots, k-1) W_{k-2}(1, \cdots, k-2) \\ &+ N_{k-1}(1, \cdots, k-1, 0) W_{k-1}(1, \cdots, k-1) \\ &+ N_k(1, \cdots, k-1, 0, k) W_k(1, \cdots, k-1, 0) \\ &\quad + \cdots + N_n(0, 1, 2, \cdots, n) W_n(0, \cdots, n-1). \quad (2.3\text{-}7) \end{aligned}$$

Using the second ordering for the linear functionals we obtain

$$
\begin{aligned}
u = N_0(1)V_0 + N_1(1,2)W_1(1) + \cdots \\
+ N_{k-2}(1,\cdots,k-1)W_{k-2}(1,\cdots,k-2) \\
+ N_{k-1}(1,\cdots,k)W_{k-1}(1,\cdots,k-1) \\
+ N_k(1,\cdots,k,0)W_k(1,\cdots,k) \\
+ \cdots + N_n(0,1,\cdots,n)W_n(0,\cdots,n-1) . \quad (2.3\text{-}8)
\end{aligned}
$$

Setting (2.3-7) equal to (2.3-8) we obtain

$$
\begin{aligned}
N_{k-1}(1,\cdots,k-1,0)W_{k-1}(1,\cdots k-1) \\
+ N_k(1,\cdots,k-1,0,k)W_k(1,\cdots,k-1,0) \\
= N_{k-1}(1,\cdots,k)W_{k-1}(1,\cdots,k-1) \\
+ N_k(1,\cdots,k,0)W_k(1,\cdots,k) .
\end{aligned}
$$

Rearranging we have

$$
\begin{aligned}
\big(N_{k-1}(1,\cdots,k-1,0) - N_{k-1}(1,\cdots,k)\big)W_{k-1}(1,\cdots,k-1) \\
= -N_k(1,\cdots,k-1,0,k)W_k(1,\cdots,k-1,0) \\
+ N_k(1,\cdots,k,0)W_k(1,\cdots,k) .
\end{aligned}
$$

By the previous remarks, this equation is equivalent to

$$
\begin{aligned}
\big(N_{k-1}(0,\cdots,k-1) - N_{k-1}(1,\cdots,k)W_{k-1}(1,\cdots,k-1) \\
= -N_k(0,\cdots,k)\; -W_k(0,\cdots,k-1 \;+ W_k(1,\cdots,k)\big) .
\end{aligned}
$$

Applying L_k to both sides and observing that

$$L_k\left(W_k(1, \cdots, k)\right) = 0 ,$$

we have

$$\left(N_{k-1}(0, \cdots, k-1) - N_{k-1}(1, \cdots, k)\right) L_k\left(W_{k-1}(1, \cdots, k-1)\right)$$

$$= - N_k(0, \cdots, k) L_k(0, \cdots, k-1) .$$

Therefore, if $L_k\left(W_k(0, \ldots, k-1)\right) \neq 0$, we have

$$N_k(0, \cdots, k) = C\left[N_{k-1}(1, \cdots, k) - N_{k-1}(0, \cdots, k-1)\right] \qquad (2.3\text{-}9)$$

where

$$C = \frac{L_k\left(W_{k-1}(1, \cdots, k-1)\right)}{L_k\left(W_k(0, \cdots, k-1)\right)} . \qquad (2.3\text{-}10)$$

We summarize the above remarks in the following theorem:

Theorem (2.3.6): Let V, $\{L_i\}$ and $\{V_i\}$ satisfy the conditions of Theorem (2.3.2). Furthermore assume $k \geq 2$ and

$$L_k\left(W_k(0, \cdots, k-1)\right) \neq 0 .$$

Then, the recursion formula is given by (2.3-9) and (2.3-10).

∎

We can now state the divided difference formula for functions of several variables.

Theorem (2.3.7): Let f be a function defined in a region G of R^n. Let $\phi_0(\bar{x}), \ldots, \phi_{n+1}(\bar{x})$ n + 2 linearly independent functions defined in G and let

and

$$E(\overline{x}) = [0, 1, 2, \cdots, x] \frac{\det \begin{bmatrix} \phi_0(\overline{x}_0) & \cdots & \phi_0(\overline{x}_n) & \phi_0 \\ \vdots & & \vdots & \vdots \\ \phi_{n+1}(\overline{x}_0) & \cdots & \phi_{n+1}(\overline{x}_n) & \phi_{n+1} \end{bmatrix}}{\det \begin{bmatrix} \phi_0(\overline{x}_0) & \cdots & \phi_0(\overline{x}_n) \\ \vdots & & \vdots \\ \phi_n(\overline{x}_0) & \cdots & \phi_n(\overline{x}_n) \end{bmatrix}} .$$

<u>Proof</u>: Let $\overline{x}_0, \ldots, \overline{x}_n$ belong to G and let $\overline{x}$ be an arbitrary point in G. The result follows from Theorem (2.3.3) by defining $L_i(\phi)$ to be the evaluation functionals

$$L_i(\phi) = \phi(\overline{x}_i) ,$$

$i = 0, 1, 2, \ldots, n$ and $L_{\overline{x}}(\phi)$ the functional

$$L_{\overline{x}} = \phi(\overline{x}) .$$

■

As in the single variable case we can compute the divided differences according to the following scheme

$$\begin{array}{ccccccc}
[0] & & & & & & \\
 & \searrow & & & & & \\
 & & [0\ 1] & & & & \\
 & \nearrow & & \searrow & & & \\
[1] & & & & [0\ 1\ 2] & & \\
 & \searrow & & \nearrow & & \searrow & \\
 & & [1,\ 2] & & & & [0,\ 1,\ 2,\ 3] \\
 & \nearrow & \vdots & \searrow & & \nearrow & \vdots \\
[2] & & \vdots & & [1\ 2\ 3] & & \vdots \\
\vdots & & \vdots & & \vdots & & \vdots
\end{array}$$

For computational purposes it is also helpful to keep track of the various determinants in formula (2.3-10) for the constant C.

Theorem (2.3.7) shows that Newton's divided difference formula for functions of several variables is considerably more complicated than the corresponding formula for the single variable case. However there are certain special cases where the formula can be considerably reduced. For divided differences on a triangular grid the reader should consult [21] p. 372. For divided differences on a rectangular grid see [19] p. 31.

2.4 AITKEN INTERPOLATION

In interpolation for a single variable we can solve for the interpolation formula according to the following scheme. If $P_{i_1,\ldots,i_p}(x)$ is the unique polynomial such that

$$P_{i_1,\ldots,i_p}\left(x_{i_j}\right) = f\left(x_{i_j}\right)$$

for $j = 1, \ldots, p$, then

$$P_{i_1,\ldots,i_p}(x) = \frac{1}{x_{i_p} - x_{i_{p-1}}} \det \begin{bmatrix} P_{i_1,\ldots,i_{p-2},i_p}(x) & x - x_{i_{p-1}} \\ P_{i_1,\ldots,i_{p-1}}(x) & x - x_{i_{p-1}} \end{bmatrix}.$$

We can compute $P_{0,1,\ldots,n}(x)$ by the following scheme

$$\begin{array}{llll}
P_0 & & & \\
P_1 & P_{0,1} & & \\
P_2 & P_{0,2} & P_{0,1,2} & \\
\vdots & \vdots & \vdots & \ddots \\
P_n & P_{0,n} & P_{0,n-1,n} \ldots & P_{0,1,2,\ldots,n}
\end{array}$$

It can be shown that the Aitken formula can be derived directly from the Lagrange formula using the properties of Vandermonde determinants. For this reason it does not appear that a scheme similar to Aitken's scheme should be valid for the several variable case. However, there is an Aitken type scheme,

for functions of several variables, which is due to Thacher and Milne [19], which is valid for a restricted class of points. In this generalization, we add enough points at each step to obtain the polynomial of next highest total degree. The scheme is valid in R^n, however for simplicity we consider all the points in R^2.

A sufficient set of conditions to guarantee a generalized Aitken method is the following.

Definition (2.4.1): A set S in R^2 is called an acceptable set if the following are satisfied

1) S^n has $(2+n)!/2!\,n!$ inhomogeneously degree n independent points

2) S^n can be factored into the subsets S_1^{n-1}, S_2^{n-1} and S_3^{n-1} such that

a) $S^n = S_1^{n-1} \cup S_2^{n-1} \cup S_3^{n-1}$

b) $S^n - \bigcup_{\substack{j=1 \\ j \neq k}}^{3} S_j^{n-1} = \{\bar{x}_k\}$ for

$$k = 1, 2, 3.$$

c) The set $B = \{\bar{x}_k\}_{k=1,2,3}$ defined by (b) is inhomogeneously degree 1 independent

d) If $\bar{x}$ belongs to S^n but not to S_k^{n-1} then $\bar{x}$ is an inhomogeneous degree 1 combination of $B - \{\bar{x}_k\}$.

e) Each set S_k^{n-1} is an acceptable set by the above definition.

Aitken's generalization takes the following form:

Theorem (2.4.2): Let S^n be a set of $(2+n)!/2!\,n!$ points satisfying the above conditions and let f be a function defined in a region G of R^2 containing the set S^n. Then, there exists a unique polynomial P_n of total degree n such that

$$P_n(x_i, y_i) = f(x_i, y_i)$$

for all (x_i, y_i) in S^n. In this case

$$P_n(x, y) = \frac{\det\begin{bmatrix} P_{n-1}^1(x, y) & (x_1 - x) & (y_1 - y) \\ P_{n-1}^2(x, y) & (x_2 - x) & (y_2 - y) \\ P_{n-1}^3(x, y) & (x_3 - x) & (y_3 - y) \end{bmatrix}}{\det\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}}, \tag{2.4-1}$$

where the set

$$B = \{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$$

is defined by condition (c) in Definition (2.4.1) and

$$P_{n-1}^k(x_i, y_i) = f(x_i, y_i)$$

for all (x_i, y_i) in S_k^{n-1} $(k = 1, 2, 3)$.

Proof: The existence of $P_n(x, y)$ follows immediately from Theorem (1.4.11). We must show formula (2.4-1) is valid. We remark that the determinant in the denominator is not zero since the points in B are inhomogeneously degrees 1 independent. Proceeding inductively, we notice since each $P_{n-1}^k(x, y)$ has total degree $n - 1$, the polynomial $P_n(x, y)$ has total degree at most n. It remains to show

$$P_n(x_i, y_i) = f(x_i, y_i)$$

for all (x_i, y_i) in S^n. We divide the proof into two cases.

Case I. ((x_i, y_i) belongs to B). Without loss of generality we choose the point (x_1, y_1). Expanding (2.4-1) we obtain

$$P_n(x_1, y_1) = \frac{P^1_{n-1}(x_1, y_1) \det \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{bmatrix}}{\det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}}. \qquad (2.4\text{-}2)$$

Since (x_1, y_1) belongs to S_1^{n-1}, we have

$$P^1_{n-1}(x_1, y_1) = f(x_1, y_1).$$

Therefore, since the determinants in Equation (2.4-2) have the same value we have

$$P_n(x_1, y_1) = f(x_1, y_1).$$

Case II. In this case we assume (x_i, y_i) belongs to S^n but is not one of the points (x_1, y_1), (x_2, y_2) and (x_3, y_3). By condition (2-b) in Definition (2.4.1) (x_i, y_i) must belong to at least two of the sets S_k^{n-1}. Without loss of generality we assume (x_i, y_i) belongs to S_1^{n-1} and S_2^{n-1} but not to S_3^{n-1}. Then

$$P_n(x_i, y_i) = \frac{\det \begin{bmatrix} f(x_i, y_i) & (x_1 - x_i) & (y_1 - y_i) \\ f(x_i, y_i) & (x_2 - x_i) & (y_2 - y_i) \\ P^3_{n-1}(x_i, y_i) & (x_3 - x_i) & (y_3 - y_i) \end{bmatrix}}{\det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}}.$$

The determinant in the numerator can be rewritten by subtracting the first row from rows two and three; we obtain

$$P_n(x_i, y_i) = \frac{\det \begin{bmatrix} f(x_i, y_i) & (x_1 - x_i) & (y_1 - y_i) \\ 0 & (x_2 - x_1) & (y_2 - y_1) \\ P^3_{n-1}(x_i, y_i) - f(x_i, y_i) & (x_3 - x_1) & (y_3 - y_1) \end{bmatrix}}{\det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}} .$$

Expanding the numerator we have

$$P_n(x_i, y_i) = f(x_i, y_i) \frac{\det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}}{\det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}}$$

$$+ P^3_{n-1}(x_i, y_i) \frac{\det \begin{bmatrix} x_1 - x_i & y_1 - y_i \\ x_2 - x_1 & y_2 - y_1 \end{bmatrix}}{\det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}} .$$

However, since

$$\det \begin{bmatrix} x_1 - x_i & y_1 - y_i \\ x_2 - x_1 & y_2 - y_1 \end{bmatrix} = \det \begin{bmatrix} 1 & x_i & y_i \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{bmatrix} ,$$

and since (x_i, y_i), (x_1, y_1) and (x_2, y_2) are inhomogeneous degree 1 dependent (condition (2-d)), we have

$$\det \begin{bmatrix} x_1 - x_i & y_1 - y_i \\ x_2 - x_i & y_2 - y_i \end{bmatrix} = 0 .$$

Therefore

$$P_n(x_i, y_i) = f(x_i, y_i) .$$

Suppose (x_i, y_i) belongs to S_1^{n-1}, S_2^{n-1} and S_3^{n-1}. Then Equation (2.4-1) gives

$$P_n(x_i, y_i) = \frac{f(x_i, y_i) \det \begin{bmatrix} 1 & x_1 - x_i & y_1 - y_i \\ 1 & x_2 - x_i & y_2 - y_i \\ 1 & x_3 - x_i & y_3 - y_i \end{bmatrix}}{\det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}}$$

$$= f(x_i, y_i) .$$

$\blacksquare$

It is difficult to prove whether or not a given set is acceptable. However, we can prove a triangular grid is acceptable set.

Theorem (2.4.3): The triangular grid in Figure (2.4.1) is an acceptable set.

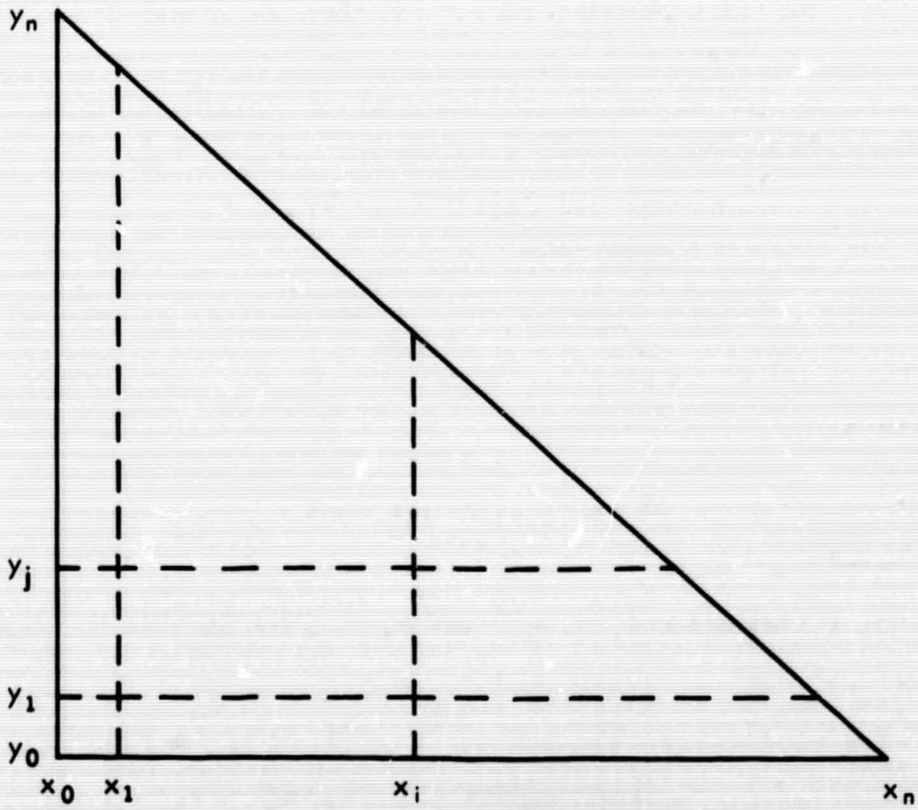

Figure (2.4.1)

Proof: We prove the theorem by induction. First let $n = 1$. We must show the set

$$S^1 = \{(x_0, y_1), (x_0, y_0), (x_1, y_0)\}$$

is acceptable. Since

$$\det \begin{bmatrix} 1 & x_0 & y_0 \\ 1 & x_0 & y_1 \\ 1 & x_1 & y_0 \end{bmatrix}$$

$$= -(x_1 - x_0)(y_1 - y_0)$$

$$\neq 0 , \qquad (2.4\text{-}3)$$

and since S^1 has

$$\frac{(2+1)!}{2!\,1!} = 3$$

points, it follows that condition (1) is satisfied. If we let

$$S_1^0 = \{(x_0, y_0)\}$$

$$S_2^0 = \{(x_0, y_1)\}$$

$$S_3^0 = \{(x_1, y_0)\} ,$$

then conditions (2-a) and (2-b) are satisfied. For this case Equation (2.4-3) shows that (2-c) is satisfied. The reader can check that (2-d) and (2-e) are easily satisfied.

Next we assume the theorem is true for the case $n - 1$ and show it holds for n. We divide the triangle into three parts as shown in Figure (2.4.2).

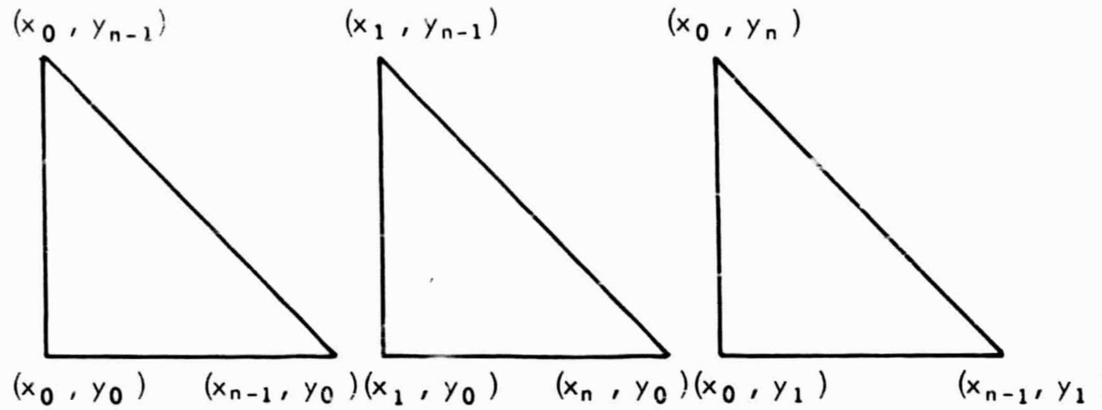

Figure (2.4.2)

Since the triangle has

$$\frac{(2+n)!}{2!\,n!}$$

inhomogeneous degree n independent points (Theorem (1.3.3) and Theorem (1.4.11)), it follows that condition (1) in Definition (2.4.1) is satisfied. Define

$$B = \left\{(x_0, y_n), (x_n, y_0), (x_0, y_0)\right\}.$$

Since

$$\det \begin{bmatrix} 1 & x_0 & y_n \\ 1 & x_0 & y_0 \\ 1 & x_n & y_0 \end{bmatrix}$$

$$= (x_n - x_0)(y_n - y_0)$$

$$\neq 0,$$

we see that B is inhomogeneously degree 1 independent. Furthermore from the definitions for B, S_1^{n-1}, S_2^{n-1} and S_3^{n-1}, it follows that conditions (2-a), (2-b), and (2-c) are satisfied. To show condition (2-d) is valid, let (x_i, y_j) belong to S^n but not S_1^{n-1}. The point (x_i, y_j) lies on the straight line between (x_0, y_n) and (x_n, y_0); therefore the set

$$\left\{(x_0, y_n) \quad (x_n, y_0) \quad (x_i, y_j)\right\}$$

is inhomogeneous degree 1 dependent. Therefore, (x_i, y_j) is an inhomogeneous degree 1 combination of $B - \{(x_0, y_0)\}$ which proves (2-d). Condition (2-e) follows from the induction hypothesis. ∎

The following theorem shows a non-singular linear transformation of an acceptable set is acceptable. For a proof we refer the reader to Thacher's paper ([18], p. 621, 622).

Theorem (2.4.4): Let T be a non-singular linear transformation. Then the sets $\{x_j\}$ and $\{Tx_j\}$ are both (homogeneously or inhomogeneously) degree n dependent or degree n independent.

REFERENCES

1. H. B. Curry, Abstract differential operations and interpolation formulas, Portugallae Mathematica 10 (1951) Fasc. 4 pp. 135-162.

2. M. Christian Coatmelec, Approximation et interpolation des fonctions differ-entiables de plusieurs variables, Ann. seient éc. norm. sup.; 3[e] série (83) 1966 pp. 271-341.

3. P. J. Davis, Interpolation and Approximation, Blaisdell, New York, 1963.

4. W. Feller, An Introduction to Probability Theory and its Applications, Vol. I, John Wiley and Sons, New York, 1950.

5. F. B. Hildebrand, Introduction to Numerical Analysis, McGraw-Hill, New York, 1956.

6. P. M. Hirsch, Numerical evaluation and estimation of multiple integrals, Ph.d. dissertation, Univ. of Wis. 1966.

7. A. S. Householder, Principles of Numerical Analysis, McGraw-Hill, New York, 1953.

8. W. E. Milne, W. Arntzen, N. Reynolds, and J. Wheelock, Mathematics for digital computers, Volume I: multivariate interpolation, WADC TR 57-556, Wright Air Development Center, Wright Patterson Air Force Base, Ohio, February 1958.

9. H. E. Salzer, Note on interpolation for a function of several variables, Bull. Amer. Math. Soc. 51 (1945) pp. 279-280.

10. H. E. Salzer, On Numerical Approximations, R. E. Langer, Ed., Univ. of Wis., Press, Madison, Wis., 1959, pp. 61-98.

11. H. E. Salzer, Divided differences for functions of two variables for irregularly spaced arguments. Numerische Matematik 6 (1964), pp. 68-77.

12. A. Sard, Remainders: Functions of several variables, Acta Math. 84 (1951), pp. 319-346.

13. A. Sard, New function spaces and their adjoints, Ann., New York Acad. Sci., 86 (1960), pp. 700-757.

14. S. L. Sobolev, On the interpolation of functions of n variables, Soviet Math., 2 (1961), pp. 343-346.

15. J. F. Steffenson, Interpolation, Baltimore, Williams and Wilkins, 1927.

16. H. C. Thacher, Jr., Generalization of concepts related to linear dependence, J. Soc. Indust. Appl. Math. 6 (1958), pp. 288-289.

17. H. C. Thacher, Jr., Derivation of interpolation formulas in several independent variables, Ann., New York Acad. Sci., 86 (1960), pp. 758-775.

18. H. C. Thacher, Jr., Numerical integration and differentiation of functions of several independent variables by an iterative procedure, J. Soc. Indust. Appl. Math., 11 (1963), pp. 614-622.

19. H. C. Thacher, Jr., and W. E. Milne, Interpolation in several variables, J. Soc. Indust. Appl. Math., 8 (1960), pp. 33-42.

20. G. Whaples, A note on degree-n independence, J. Soc. Indust. Appl. Math., 6 (1958), pp. 300-301.

21. E. T. Whittaker and G. Robinson, Calculus of Observations, 4th ed., London, Blackie and Son, 1944.

22. K. Hoffman and R. Kunze, Linear Algebra, Prentice Hall, Englewood Cliffs, New Jersey, 1961.